Search of new resonances decaying into top quark pairs with the ATLAS detector at the LHC and jet calibration studies

Reina Camacho

DU 2259
EDSF 719

PCCF T 1203

UNIVERSITÉ BLAISE PASCAL DE CLERMONT-FERRAND

U.F.R. Sciences et Technologies

ÉCOLE DOCTORALE DES SCIENCES FONDAMENTALES

T H È S E

pour obtenir le grade de

Docteur d'Université

Specialité : Physique des Particules

Présentée et soutenue par

Reina Camacho

Recherche de nouvelles résonances se désintégrant en paires de quarks top avec le détecteur ATLAS du LHC

Thèse soutenue publiquement le 13 Juillet 2012 devant la commision d'examen

Président :	Alain Falvard	-	CNRS (LPC Clermont-Ferrand)
Rapporteurs :	Aurelio Juste Rozas	-	Universidad Autonoma de Barcelona (IFAE)
	Gustaaf Brooijmans	-	Columbia University
Examinateurs :	Yves Sirois	-	CNRS (Laboratoire Leprince Ringuet)
	Tancredi Carli	-	CERN
	Henry Bachacou	-	CNRS (CEA-Saclay)
Invités :	Emmanuel Busato	-	Université Blaise Pascal-LPC Clermont-Ferrand
	Samuel Calvet	-	CNRS (LPC Clermont-Ferrand)
Directeur de thèse :	Dominique Pallin	-	CNRS (LPC Clermont-Ferrand)

DU 2259
EDSF 719

PCCF T 1203

UNIVERSITÉ BLAISE PASCAL OF CLERMONT-FERRAND

U.F.R. Sciences et Technologies

ÉCOLE DOCTORALE DES SCIENCES FONDAMENTALES

T H E S I S

to obtain the title of

PhD of Science

Specialty : Particle Physics

Defended by

Reina Camacho

Search of new resonances decaying into top quark pairs with the ATLAS detector at the LHC and jet calibration studies

Defended on July 13^{th}, 2012

President :	Alain Falvard	-	CNRS (LPC Clermont-Ferrand)
Reviewers :	Aurelio Juste Rozas	-	Universidad Autonoma de Barcelona (IFAE)
	Gustaaf Brooijmans	-	Columbia University
Examinators :	Yves Sirois	-	CNRS (Laboratoire Leprince Ringuet)
	Tancredi Carli	-	CERN
	Henry Bachacou	-	CNRS (CEA-Saclay)
Invited :	Emmanuel Busato	-	Université Blaise Pascal-LPC Clermont-Ferrand
	Samuel Calvet	-	CNRS (LPC Clermont-Ferrand)
Advisor :	Dominique Pallin	-	CNRS (LPC Clermont-Ferrand)

Abstract

Theoretical context

The Standard Model (SM) of particle physics summarizes the current knowledge of subatomic physics and the known fundamental interactions (except for the gravitation). Up to now there is no experimental result that strongly contradicts the SM predictions. However, there are still some missing pieces in the SM of experimental or theoretical nature. Several physical models, known as Beyond-the-SM (BSM) theories, has been proposed in the last decades to try to account for one or more of the SM's open questions. Some of these theories as the topcolor model, chiral color models and Randall-Sundrum models with warped extra dimensionas predict the existence of new particles that couple strongly to the top quark, due to its high mass. This coupling implies that they decay primarily to top-quark pairs, $t\bar{t}$.

To test the validity of the SM and/or of the BSM theories, high energy accelerators of the TeV order are needed, as well as detectors adapted to analyse the collected events. Therefore, several searches for $t\bar{t}$ resonances have been performed at hadron-hadron colliders as the Tevatron or the Large Hadron Collider (LHC). The benchmark models considered are a topcolor assisted technicolor (TC2) which produces a topphilic Z' particle (used as benchmark for narrow $t\bar{t}$ resonances) and a Randall-Sundrum (RS) warped extra-dimension that would result in a bulk Kaluza-Klein (KK) gluon (used as benchmark for wide $t\bar{t}$ resonances). The topcolor assisted technicolor model explain the top mass and electroweak symmetry breaking through top quark condensation associated with symmetry breaking of a new strong force. The Randall-Sundrum model with a single extra warped extra dimension explain the hierarchy problem by allowing the gravity to propagate in the added warped dimension. The CDF and DØ collaborations have excluded a leptophobic topcolor Z' with mass smaller than 900 GeV at 95% C.L. At the LHC, the best limit obtained for the Z' and the KK gluon has been obtained by the CMS experiment using 2011 data, $500 < m_{Z'} < 1300$ GeV and $1000 < m_{g_{KK}} < 1400$ GeV, respectively.

Experimental context

The studies presented in this thesis were performed using data collected by the ATLAS detector at the LHC. The LHC is the largest and highest energy particle accelerator ever built. Around 10000 physicists and engineers around the world are taking part in this experience by developing new techniques and approaches to identify the interesting physics buried in the complex environement produced in the LHC pp collisions. The LHC produced its first pp collisions on November 23, 2009 at the injection energy of 450 GeV. On March 19, 2010 the LHC broke a record by raising the beam energy to 3.5 TeV, and the first pp collisions at 7 TeV were recorded on March 30, 2010. The beam energy for the whole 2011 year was 3.5 TeV per beam, while in 2012 the beam energy is 4 TeV. At the beginning of 2013 the LHC will go into a long shutdown to prepare for higher energy collisions starting in 2014. At the four collision points of the LHC, detectors have been placed to study the high-energy collisions. One of these detectors, ATLAS, a general purpose detector with an extensive initial physics program. The ATLAS detector consists of a tracking system in a 2 T solenoid field, providing coverage up to a pseudo-rapidity of $|\eta| < 2.5$, sampling electromagnetic and hadronic calorimeters up to $|\eta| < 4.9$, muon chambers in a toroidal magnetic field and a trigger system consisting of three levels of event selection. So far, the ATLAS sub-detectors have shown an excellent performace in terms of efficiency and resolution.

This thesis contains three different analyses interconnected. The main one concerns the results of the search for new resonances that decay to top-quark pairs using the first 2.05 fb^{-1} of data collected by the ATLAS detector in 2011. Secondly and related to this search, performance studies of the Jet Vertex Fraction (JVF) in top-quark pairs topologies are presented too. JVF is a variable that can be used to reduce the pile-up effects to improve the precision and sensitivity of physics analyses at high luminosities. Finally, results regarding the performance, validation in data and associated systematic uncertainty derivation of the Global Sequential (GS) jet calibration are discussed. The determination of the jet energy scale and the achievement of the optimal jet performance is of key importance to many LHC physics analyses, specially to the main analysis of this thesis due to the presence of jets in the final state. The results are presented in order that they were performed during the thesis.

Global Sequential jet calibration

During the first years of data analysis in ATLAS the jet energy scale was obtained using a simple calibration approach, known as EM+JES calibration. EM+JES is a simple calibration scheme but with a low performance regarding the jet energy and angular resolution and the flavor sensitivity. The Global Sequential (GS) calibration is a multivariate extension of the EM+JES calibration derived in simulated events. It consists in removing the dependence of the calorimeter response on several transverse and longitudinal jet properties. These corrections improve the jet energy resolution and reduce the jet response flavour sensitivity, without changing the jet energy scale. The longitudinal structure of the jet is characterized by the fraction of jet energy deposited in the different layers of the ATLAS calorimeter, while its transverse structure is characterized by the jet width. The key features of the GS are:

- Good performance. The GS jet resolution improvement with respect to EM+JES in data and simulation is comparable to the one achieved by other sophisticated jet energy calibration schemes developed in ATLAS and goes up to 30% for jets with a transverse momentum of 200 GeV. On the other hand, the GS jet response flavour sensitivity (estimated throught the difference between the light quark-jet and gluon-jet average response) improve by 3% (2%) at low (high) p_{T} with respect to EM+JES.

- Its performance is not very sensitive to pile-up (in the pile-up conditions in 2010).

- The GS corrections were validated in data directly using a method known as *Di-jet Balance* (about 35 pb^{-1} of data collected by the ATLAS experiment in 2010 were used). The results show a good agreement between data and simulation.

- The derivation of the GS associated systematic uncertainty is easy. It is estimated in an inclusive multi-jets sample using a method based on the differences between data and Monte Carlo of the jet properties used as input to the calibration. The latter is 0.5% for $30 < p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV and $|\eta| < 2.1$ and 1% for $p_{\mathrm{T}}^{\mathrm{jet}} < 30$ GeV and $2.1 < |\eta| < 2.8$. These results are supported by the results obtained on the γ+jets events in a p_{T} range between 20 GeV and 260 GeV. This uncertainty has to be added in quadrature to the EM+JES calibration systematic uncertainty. Extra systematic uncertainties need to be added to account for the dependence of the jet response on the event topology and jet flavour in the future.

The results of this work can be found in the following Refs. [1, 2, 3].

Search for $t\bar{t}$ resonances in ATLAS in the lepton plus jets channel

A search for $t\bar{t}$ resonances using 2.05 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV collected in 2011 with ATLAS was performed. The search is made in the lepton plus jets channel where one W from the top decays leptonically (to an electron or a muon plus the corresponding neutrino) and the other W hadronically (into a quark-antiquark pair). To find a $t\bar{t}$ resonance a serie of cuts are applied to enhance the $t\bar{t}$ topology of the final state and reduce the background. The selection used in the analysis is designed to account for the $t\bar{t}$ resolved (all the objects in the final state are individually identified) and semi-boosted (two of the partons from the hadronic top decay are reconstructed as a single object) topologies. The $t\bar{t}$ candidate invariant mass was reconstructed after requiring an electron or a muon, substantial missing transverse energy and at least four jets with high p_{T}, of which at least one was tagged as a jet coming from a b-quark. Four different methods to reconstruct the $t\bar{t}$ system have been tested. Some of them shown a clear dependence on the pile-up level. Finally, the chosen method use jets that are "close" to the rest of the activity in the event, rejecting in this way jets coming from initial state radiation. The biggest backgrounds for this analysis are: SM $t\bar{t}$, W+jets, single top and multijet events with misidentified leptons.

The reconstructed $t\bar{t}$ invariant mass is found to be compatible with the SM prediction. As no evidence of a resonance has been found 95% C.L. upper limits are set on the production cross-section times branching ratio to top quark pairs for a narrow Z' and a wide Randall-Sundrum Kaluza-Klein gluon resonances. For the Z' resonance the 95% C.L. observed upper limit range from 9.3 pb for a mass of 500 GeV to 0.95 pb for a mass of 1300 GeV, excluding a leptophobic topcolor Z' boson with 500 GeV$< m_{Z'} <$880 GeV. For the KK gluons the 95% C.L. observed upper limit range from 10.1 pb for a mass of 500 GeV to 1.6 pb for a mass of 1300 GeV. KK gluons in the RS model with masses between 500 and 1130 GeV are excluded at 95% C.L. These results are part of the paper in Ref. [4].

Jet Vertex Fraction

The Jet Vertex Fraction (JVF) is a variable calculated for each jet that quantifies the fraction of track transverse momentum associated to a jet from the hard-scattering interaction. An event selection using the JVF variable can be effectively used to distinguish jets coming from pile-up interactions from jets coming from the hard-scatter interaction. The optimal JVF cut for top-quark analyses in 2011 was found to be $|JVF| > 0.75$. Scale factors to account for differences in the performance of the JVF cut in data and simulation were calculated using a tag& probe method developed using $Z(\rightarrow ee/\mu\mu)$+jets events. These scale factors were shown to largely recover the agreement between data and simulation in the $t\bar{t}$ analysis in the lepton plus jets channel. The systematic uncertainty associated to these scale factor were also calculated. Variations of the order of 2.5-3% were observed for the systematic uncertainty. These results are being used in most of the top analyses with the whole dataset collected in 2011, corresponding to an integrated luminosity of 4.7 fb^{-1} [5, 6].

Keywords: ATLAS, LHC, top quark, jet calibration, Global Sequential calibration, resonances top-antitop, Kaluza-Klein gluon, leptophobic topcolor Z', Jet Vertex Fraction, pile-up.

Résumé

Contexte théorique

Le Modèle Standard (MS) de la physique des particules synthétise les connaissances actuelles sur la physique subatomique et les interactions fondamentales connues (sauf la gravitation). Jusqu'à présent, aucun résultat expérimental ne contredit fortement les prédictions du MS. Cependant, il existe encore quelques questions de nature expérimentale ou théorique qui restent sans réponse dans le MS. De nombreux modèles théoriques se proposant de décrire la physique au delà du MS ont été proposés dans les dernières décennies pour tenter de remédier à une ou plusieurs des lacunes du MS. Certaines de ces théories comme le modèle technicouleur, les modèles de couleurs chiraux et les modèles de Randall-Sundrum avec des dimensions supplémentaires prédisent de nouvelles particules qui se couplent fortement au quark top, en raison de sa masse élevée. Ce couplage implique une désintégration préférentielle en une paire de quarks top, $t\bar{t}$.

Pour pouvoir tester ces nouvelles théories, des accélérateurs capables d'atteindre des énergies de l'ordre du TeV sont nécessaires, de même que des détecteurs adaptés pour analyser les événements recueillis. Ainsi, plusieurs recherches de résonances top-antitop ont été réalisées auprès des collisionneurs hadroniques très puissants comme Tevatron et le grand collisionneur de hadrons (LHC: acronyme de *Large Hadron Collider*). Les modèles de référence considérés dans ces recherches sont le modèle *topcolor assisted technicolor* qui conduit à la production d'un boson neutre Z' et le modèle de Randall-Sundrum avec une dimension supplémentaire qui conduit à la production d'un gluon Kaluza-Klein qui se désintégre aussi en une paire $t\bar{t}$. Le modèle *topcolor assisted technicolor* explique la masse élevée du quark top et la brisure de symétrie électrofaible par condensation du quark top associé à la brisure de symétrie d'une nouvelle force forte. Le modèle de Randall-Sundrum explique le problème de hiérarchie en prenant en compte une dimension supplémentaire enroulée sur elle-même et dans laquelle la gravité va se propager. Les expériences CDF et DØ ont exclu à 95% de C.L. une masse de Z' inférieure à 900 GeV. Au LHC, la meilleures limites sur les masses du boson Z' et du gluon de Kaluza Klein ont été obtenues par l'expérience CMS en utilisant les données collectées en 2011, $500 < m_{Z'} < 1300$ GeV et $1000 < m_{g_{KK}} < 1400$ GeV, respectivement.

Contexte expérimental

Le travail de recherche exposé dans cette thèse a été réalisé en utilisant les données collectées par le détecteur ATLAS auprès du LHC. Le LHC est l'accélérateur de particules le plus grand et le plus puissant jamais construit. Il est constitué d'un double anneau de stockage de protons. Environ 10000 physiciens et ingénieurs du monde entier participent à cette expérience en développant de nouvelles techniques et approches pour identifier les événements intéressants cachés dans l'environement complexe produit dans des collisions proton-proton. Les premières collisions proton-proton ont eu lieu à la fin de l'année 2009 à une énergie de 900 GeV dans le système du centre de masse. Le 19 Mars 2010, le LHC a battu un record en augmentant l'énergie des faisceaux à 3.5 TeV, et les premières collisions à 7 TeV ont été enregistrées le 30 Mars 2010. L'énergie par faisceau pour l'ensemble de 2011 était de 3.5 TeV, et 4 TeV en 2012. Un an d'arrêt technique est prévu en 2013 avant de parvenir à l'énergie nominale de collision de 14 TeV. Quatre expériences se partagent les quatre points de croisement des faisceaux du grand anneau du LHC. Une d'entre elles, ATLAS, est un détecteur généraliste avec un vaste programme de physique.

ATLAS est constitué d'un détecteur interne de traces dans un champ magnétique de 2 T, offrant une couverture jusqu'à $|\eta| < 2.5$, un système calorimétrique allant jusqu'à $|\eta| < 4.9$, un spectromètre à muons dans un champ magnétique toroïdal et un système de déclenchement composé de trois niveaux. Tous les sous-systèmes ont d'excellentes performances en termes d'efficacité et de résolution.

Cette thèse comprend trois travaux interconnectés. En premier lieu, elle décrit les résultats de la recherche de nouvelles résonances qui se désintègrent en paires $t\bar{t}$ en utilisant les premiers 2.05 fb^{-1} de données collectées par le détecteur ATLAS en 2011. Dans une deuxième partie, en connection avec cette recherche, des études de performance de *Jet Vertex Fraction* (JVF) en utilisant des événements $t\bar{t}$ sont également présentées. JVF est une variable qui peut être utilisée pour réduire les effets d'empilement afin d'améliorer la précision et la sensibilité des analyses de physique à haute luminosité. Finalement, les performances de la calibration Globale Séquentielle des jets, sa validation sur des données réelles et l'évaluation de l'incertitude systématique qui lui est associée seront aussi discutées. La détermination précise de l'échelle en énergie des jets (JES: acronyme de *Jet Energy Scale*) ainsi que l'obtention d'une résolution optimale sont extrêmement importantes pour de nombreuses analyses de physique au LHC. Cela est vrai en particulier pour la recherche de nouvelles résonances $t\bar{t}$ en raison de la présence de jets dans l'état final. Les résultats sont présentés dans l'ordre dans lequel ils ont été réalisés au cours de cette thèse.

Calibration Globale Séquentielle des Jets

Pendant les premières années d'analyse de données dans ATLAS, l'échelle en énergie des jets a été obtenue en utilisant une approche connue sous le nom calibration EM+JES. La calibration EM+JES offre une calibration simple dans son principe mais sa résolution angulaire et en énergie, ainsi que sa sensibilité à la saveur ne sont pas optimales. La calibration Globale Séquentielle (GS) est une extension multivariée de la calibration EM+JES calculée avec des événements simulés. Elle consiste à supprimer la dépendance de la réponse du calorimètre à plusieurs propriétés du jet caractérisant sa structure longitudinale et transversale. La résolution en énergie des jets se trouve améliorée et la sensibilité de la réponse à la saveur du jet réduite, tandis que la valeur moyenne de la réponse reste inchangée. La structure longitudinale du jet est caractérisée par la fraction d'énergie déposée dans les différentes couches longitudinales des calorimètres. La structure transversale est caractérisée par la largeur du jet. Les caractéristiques clés de la GS sont:

- Bonne performance. L'amélioration relative par rapport à EM+JES est d'environ 30% pour des jets avec une impulsion transverse de 200 GeV et la performance de la calibration GS est comparable à celles d'autres méthodes sophistiquées de calibration en énergie pour les jets developpées dans ATLAS. L'amélioration de la résolution sur les données est compatible avec celle prédite par la simulation Monte Carlo. De plus, la sensibilité à la saveur du jet (estimée par la différence entre la valeur moyenne de la réponse pour des jets venant des quarks et celle des jets venant des gluons) est réduite par rapport à EM+JES d'environ 3% (2%) à bas (haut) p_{T}.

- Sa performance n'est pas très sensible à la présence d'empilement (cela a été testé dans les conditions d'empilement en 2010).

- Les corrections GS ont été validées avec des données en utilisant la méthode dite *Di-jet Balance* (environ 35 pb^{-1} de données collectées par l'expérience ATLAS en 2010 ont été utilisées). Les résultats obtenus jusqu'à présent montrent un bon accord entre les données réelles et la simulation.

- Le calcul de l'incertitude systématique associée à GS est simple. Elle est estimée en utilisant un échantillon multi-jet inclusif, grâce à une méthode basée sur les différences des propriétés des jets entre les données et la simulation Monte Carlo, les propriétés considerées étant celles qui sont utilisées dans GS. L'incertitude systématique est égale à 0.5% pour $30 < p_{\mathrm{T}}^{\mathrm{jet}} <$ 800 GeV et $|\eta| < 2.1$ et 1% pour $p_{\mathrm{T}}^{\mathrm{jet}} < 30$ GeV et $2.1 < |\eta| < 2.8$. Ces résultats sont corroborés par des résultats obtenus en utilisant des événements γ+jets pour $20 < p_{\mathrm{T}} <$ 260 GeV. Cette incertitude doit être ajoutée en quadrature à l'incertitude systématique de EM+JES. Des incertitudes systématiques supplémentaires devront être ajoutées pour tenir compte de la dépendance de la réponse du jet selon la topologie des événements ou la saveur du jet.

Les résultats de ce travail peuvent être trouvés dans les références [1, 2, 3].

Recherche de résonances $t\bar{t}$ avec ATLAS dans le canal lepton plus jets

Une recherche de résonances $t\bar{t}$ en utilisant 2.05 fb^{-1} de collisions proton-proton à $\sqrt{s} = 7$ TeV collectées en 2011 avec ATLAS a été réalisée. La recherche se fait dans le canal lepton plus jets où un boson W venant d'un quark top se desintègre leptoniquement (en un électron ou en un muon et le neutrino correspondant) et l'autre W de façon hadronique (en une paire quark-antiquark). Pour trouver une résonance $t\bar{t}$, une série de sélections est appliquée afin de rendre prédominant le signal recherché devant les bruits de fond et renforcer la topologie de l'état final $t\bar{t}$. La sélection des événements utilisée a été conçue pour tenir compte de la topologie $t\bar{t}$ résolue (tous les objets dans l'état final sont identifiés individuellement) et la topologie $t\bar{t}$ semi-boostée (deux des partons de la désintégration du quark top hadronique sont reconstruits comme un objet unique). La masse invariante du système $t\bar{t}$ a été reconstruite après avoir requis un électron ou un muon, de l'énergie transverse manquante substantielle et au moins quatre jets de haut p_{T}, dont au moins un a été étiqueté comme un jet de b. Quatre méthodes de reconstruction du système $t\bar{t}$ ont été testées. Certaines d'entre elles montrent une dépendance claire avec le niveau d'empilement. Finalement, la méthode choisie utilise des jets qui sont proches du reste de l'activité dans l'événement, rejetant de cette façon les jets provenant de radiations dans l'état initial. Les bruits de fond les plus importants dans cette analyse sont: $t\bar{t}$ MS, W+jets, top célibataire et des événements multi-jets.

Aucun signe de nouvelle physique n'a été trouvé et des limites supérieures à 95% CL ont été posées sur le produit de la section efficace par le rapport de branchement en paires de quarks top pour un resonance Z' et un gluon de Kaluza-Klein. Pour la résonance Z' la limite supérieure observée est comprise entre 9.3 pb pour une masse de 500 GeV et 0.95 pb pour une masse de 1300 GeV. Le boson Z' avec une masse entre 500 GeV$< m_{Z'} <$880 GeV a été exclu. De même, le gluon de Kaluza-Klein avec une masse entre 500 et 1130 GeV a été exclu. Pour le gluon de Kaluza-Klein la limite supérieure observée est comprise entre 10.1 pb pour une masse de 500 GeV et 1.6 pb pour une masse de 1300 GeV. Ces résultats sont présentés dans la référence [4].

Jet Vertex Fraction

La *Jet Vertex Fraction* (JVF) est une variable calculée, pour chaque jet, qui quantifie la fraction d'impulsion transverse des traces associées à un jet provenant de l'interaction principale. Une sélection utilisant la variable JVF peut être efficacement utilisée pour distinguer les jets provenant d'interactions d'empilement des jets provenant de l'interaction principale. La coupure optimale

pour des analyses top en 2011 est measurée être $|JVF| > 0.75$. Les facteurs d'échelle pour tenir compte des différences dans les résultats de la coupure JVF entre les données et la simulation Monte Carlo ont été calculés en développant une méthode *tag & probe* et des événements $Z(\rightarrow ee/\mu\mu)$+jets. Ces facteurs d'échelle permettent de retrouver l'accord entre les données et la simulation dans l'analyse $t\bar{t}$ dans le canal lepton plus jets. L'incertitude systématique associée à ces facteurs d'échelle a également été calculée. Des variations de l'ordre de 2.5-3% ont été observées pour l'incertitude systématique. Ces résultats sont utilisés dans la plupart des analyses top avec l'ensemble des données recueillies en 2011, correspondant à une luminosité intégrée de 4.7 fb^{-1} [5, 6].

Mots clés: ATLAS, LHC, quark top, calibration des jets, Calibration Globale Séquentielle, résonances top-antitop, gluon Kaluza-Klein, topcolor Z', Jet Vertex Fraction, empilement.

Acknowledgments

I would like to thank Alain Baldit and Alain Farvard for hosting me at the LPC-Clemont Ferrand during these three years of thesis. I am specially grateful with Aurelio Juste Rozas and Gustaaf Brooijmans for accepting to be the referees of this document, and to Yves Sirois, Tancredi Carli and Henri Bachacou for accepting to be part of the jury.

I would like to thank to the Conseil régional d'Auvergne for its financial support during my three years of PhD studies. I want also thank to the HELEN program for giving me the opportunity of discover the experimental high energy physics through its intership program.

Research in the field of particle physics is done in collaboration and the work presented in this thesis is not the exception. Therefore, I would like to thanks the whole ATLAS collaboration for their efforts regarding the detector construction and monitoring and the understanding of the data used in this thesis. Certain members of the collaboration deserve special recognition from my part. First, I want to thank my advisor Dominique Pallin for receiving me in the ATLAS-LPC group, for his support when needed and for giving me the freedom to pursue research of my own chosing. My deepest gratitude goes to Emmanuel Busato and Samuel Calvet for all that they could teach me in these three years, their encouragement, guidance, availability and trust. Emmanuel, your enthusiasm for physics have been inspirational, thanks for your constructive feedback and for having the right word whenever I needed. Samuel, thanks for answering all my questions at all times, for helping me to see the big picture and for all the motivating "fais comme tu veux, c'est ta thèse". Jose Ocariz, Luis Alejandro Perez, Johan Lundberg and Andrea Messina also deserve my gratitude for their patience and guidance through my first steps in the high energy particle physics world.

Thanks to the past and present members of the ATLAS-LPC group: David, Emmanuel, Samuel, Dominique, Diane, Claudio, Nabil, François, Djamel, Emmanuelle, Loïc, Geoffrey, Daniela, Fabrice, Frédérique, Hongbo, Julien, Philippe, Renato and Christophe; for their scientific, technical and practical advice along these three years. In particular to David Calvet for his guidance through my first months at Clermont and for answering all questions about programming, to Diane Cinca for sharing her own experience as a PhD student at each step of the road and to Loïc Valéry for the continuous supply of mini-cakes and for keeping a cheerful atmosphere in the office at all times. I would like to thanks the administrative, informatic and technical services teams of the LPC for their help. In particular to Michel Crouau, Romeo Bonnefoy and Fabrice Podlyski for their inestimable expertise in evaluating the photomultipliers for the ATLAS hadronic calorimeter.

The analyses presented in this thesis would not have been possible without the support, expertise and insightful remarks of several people. Unfortunately, naming of all them is not possible. Emmanuel Busato, David Lopez Mateos, Ariel Schwartzman and Tancredi Carli, who introduced me to the world of jets and their calibration. Samuel Calvet, Gustaaf Brooijmans, Thorsten Kuhl, Tatjana Lenz, Lucia Masetti, Marcello Barisonzi, Elin Bergeaas Kuutmann and others members of the $t\bar{t}$ semileptonic resonance group for their crucial ideas and feedback on the analysis. To Weina Ji and Luz Stella Gomez for the long days passed cross-checking the cut flows. Samuel Calvet, Caterina Doglioni, Alison Alister, Giuseppe Salamanna, Ariel Schwartzman and Tancredi Carli for very useful discussions about several aspects of the JVF performance studies.

On a more personal note, I want to thank the people with who I shared my excitement at joyful times and also support me in the difficult ones. To Luis Nunez and Alejandra Melfo for their personal advise when I was deciding my future. To my friends in the LPC, in particular to

Diego Roa, Marwa Jahjah, Daniela Paredes and Patricio Ceniceros Montero for their patience whenever they had to listen about jets, resonances, scale factors, etc over lunch, a coffee or a beer. To Oliver Hach for making me discover the beautiful views of Clermont-Ferrand from the top of some volcanoes in the Chaîne des Puys. To my friends at Geneva, for trying to get me out of the office when the days were too long at CERN. In particular to Barbara Millan for many dinners, all the suggested work-mood songs and for giving me objectivity when doubts assaulted me and to Glenn Vanbavinckhove for all the catch up lunches and advice when needed. To Stefan Gadatsch for all the skype support during the writing phase.

To my Venezuelan HEP friends: Camila, Barbara, Homero, Arturo, Hebert, Joany, Daniela, Luis and Henso. Thanks for your sincere frienship and support. To Arturo Sanchez for all the good memories accumulated during the last 9 years. Camila Rangel, Rebeca Ribeiro and Anais Moller deserve my gratitude for being reliable friends and traveller companions, for all the pep talks and the phone/skype/SMS/hangouts/whapsapp support. I want to thanks Ingrid Mendoza, Rosanna Rangel and Cesar Zapata for keeping close in the distance.

Finally, I want to thank my family for all the support along this road, even if that meant been far away from home. Specially, to my mother and my father for encouraging my curiosity, for believing in me and for their continuous effort to support my dreams. This thesis is dedicated to them. Podría escribir otras 200 páginas sobre todas las cosas que les agradezco y lo orgullosa que me siento de ustedes, mis viejitos, es por eso que esta tesis esta dedicada a ustedes.

Contents

Introduction

The Standard Model (SM) of particle physics, built over the last half century, provides a nearly complete picture of the known particles and the way they interact with each other. However, there are still some missing pieces in the SM of experimental and theoretical order. Several physical models, called Beyond-the-SM (BSM) theories, has been proposed in the last years to try to account for one or more of the SM's open questions. The validity of the SM and/or of the BSM theories can be tested at high energy experiments and is at this point that the Large Hadron Collider (LHC) experiment enters in the scene. The LHC is the largest and highest energy particle accelerator ever built. Around 10000 physicists and engineers around the world are taking part in this experience by developing new techniques and approaches to identify the interesting physics buried in the complex environement produced in the LHC pp collisions. At the four collision points of the LHC, detectors have been placed to study the high-energy collisions. One of these detectors is ATLAS, a general purpose detector with an extensive initial physics program that includes precision measurements in the SM frame, the search of the Higgs boson and the search of signatures of new physics. An example of new physics signatures would be the existence of a new heavy particle that decays into top-quark pairs, a top pair resonance. Several BSM theories predicts this kind of heavy resonances that strongly couples to top quarks due to its high mass.

This thesis presents the results of the search for new resonances that decay to top-quark pairs in the lepton plus jets final state using the first 2.05 fb^{-1} of data collected in 2011 by the ATLAS detector. Related to this search, performance studies of the Jet Vertex Fraction (JVF) in top-quark pairs topologies are presented too. JVF is a variable that can be used to reduce the pile-up effects to improve the precision and sensitivity of physics analyses at high luminosities. The lepton plus jets final state is constituted by six individually identified decay products: four jets, an electron or muon, and a neutrino. The understanding of the jet calibration has an important role in this analysis due to the presence of jets in the final state. During the first two years of data analysis in ATLAS the jet energy scale was obtained using a simple calibration approach, easy to understand and easy to derive systematic uncertainty for it, but with a low performance regarding the jet energy and angular resolution and the flavor sensitivity. An extension of this simple calibration called Global Sequential (GS) calibration has been proposed in ATLAS. GS uses several jet properties to improve the jet energy resolution and reduce the jet response flavour sensitivity. The performance, the data validation and the associated systematic uncertainty derivation of the GS scheme are also presented in this thesis.

This thesis uses the so-called natural units, a convention commonly employed by the experimental high energy physics community where $c = h/2\pi = 1$, where c is the speed of light and h is the Planck's constant. Therefore, masses and momenta are given in GeV rather than GeV/c^2 and GeV/c, respectively.

Outline

After a brief introduction on the theoretical framework of the SM of particle physics and some BSM models in Chapter 1, the LHC complex at CERN and ATLAS detector are described in Chapter 2. Chapter 3 describes the event simulation process in ATLAS and the Monte Carlo generators used to produce the simulated samples used along this thesis. Chapter 4 describes the reconstruction, identification and calibration of physics objects combining the information

of the ATLAS sub-detectors. Only those physics objects particularly relevant for this thesis are described: tracks, primary vertices, electrons, muons, neutrinos and jets. Chapter 5 is devoted to the GS calibration. Results on its performance, the validation of the calibration using data and the evaluation of the associated systematic uncertainty are presented. Chapter 6 describes the search for $t\bar{t}$ resonances in the lepton plus jets channel in 2.05 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV collected by ATLAS during March and August 2011. Details of the selection, the $t\bar{t}$ pair reconstruction methods, the systematic uncertainties affecting the analysis and the limit setting results are given. Finally, Chapter 7 includes JVF performance studies in $t\bar{t}$ topologies in simulation and data using in 4.7 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV collected by ATLAS during 2011. JVF is currently used in the $t\bar{t}$ resonances search analysis using the 4.7 fb^{-1} of data collected in 2011. This analysis is still underway, therefore no final results can be presented in this document.

Personal contributions

Research in experimental high energy physics relies heavily on collaboration. Therefore, the work presented in this thesis relies on the results obtained by many people in different analysis subgroups. Leading contributions by the author to the various analyses presented in this thesis, divided by chapter, are presented below.

Chapter 5: GS calibration The author derived the GS calibration constants in Monte Carlo for two software releases in 2010. She performed jet resolution, linearity and pile-up sensitivity studies in Monte Carlo, as well as the validation of the GS using data collected in 2010 and the evaluation of the associated systematic uncertainty using inclusive multi-jets events. The author's work can be found in the following Refs. [1, 2, 3].

Chapter 6 Search for $t\bar{t}$ resonances The author participated specifically in the implementation and running of the selection and providing the $t\bar{t}$ pair mass spectra for the nominal selection and its variations for systematics uncertainties needed for the limits setting. She also performed performance studies of the different $t\bar{t}$ pair reconstruction described in the chapter, regarding the resolution, the efficiency and the pile-up dependence. The two different variations of the χ^2 methods were developed by the author. Contributions from the author appears in Ref. [7, 8].

Chapter 7 JVF performance studies The author determined the optimal JVF threshold to use in top-quark analyses. She also has performed efficiency studies in data and Monte Carlo, derived the corresponding scale factors to match the efficiency in Monte Carlo to the one in data and calculated their associated systematic uncertainties. These results are being used in most of the top analyses with the whole dataset collected in 2011, corresponding to an integrated luminosity of 4.7 fb^{-1} [5, 6].

In addition, the author has participated as Tile shifter in the control room, and data quality shifter for the Tile calorimeter. She also performed quality test on the 400 photomultipliers bought by ATLAS as spares to replace damaged photomultipliers in the Tile calorimeter.

> *It is wrong to think that the task of physics is to find out how nature is. Physics concerns what we can say about nature.* Niels Bohr, 1927

Chapter 1

Theoretical context

1.1 Overview

The Standard Model (SM) of particle physics is a relativistic quantum field theory that attempts to describe the known fundamental particles and their interactions, except the gravitation. The SM was formulated between the 1960s and 1970s. Since then, it has been tested by several experiments and no significant deviation from its predictions have been observed (latest results of the Standard Model fit to electroweak precision data can be found in Ref. [9]). The first section of this chapter summarizes the basic concepts of the SM. Section 1.3 discusses the known limitations of the SM and some of the new physics models trying to account for SM's open questions. Some of these beyond-the-SM-theories predict the existence of new heavy particles which couple preferably to top quarks. Therefore, the last section is devoted to the description of the top quark properties. Unless stated otherwise the bibliographic references used in this chapter are [10, 11, 12].

1.2 The Standard Model

1.2.1 Elementary particles

The known fundamental particles can be divided into two classes: the *fermions* and *bosons*. Fermions have half-integer spin and obey the Fermi-Dirac statistic, while bosons have integer spin and obey the Bose-Einstein statistic. For each particle, there exists the corresponding antiparticle with identical mass spin but electric charge of opposite sign. The antiparticles are denoted with the same symbol used for the particles but with a bar added over it or by inverting the sign of the electrical sign.

Fermions are subdivided in leptons and quarks as shown in Table 1.1. There are six quarks (known as quark flavours: up, down, charm, strange, top, bottom) and six leptons (electron, electron neutrino, muon, muon neutrino, tau, tau neutrino). They are usually arranged into three generations. Corresponding fermions in each generation have similar physical behavior (as identical electric charge and isospin) but different masses. At present there is no explanation for this fermions families distribution. Each generation consists of two quarks, one electrically charged lepton and an electrically neutral lepton (neutrino). The first generation charged particles build all stable matter. The second and third generation charged particles have only be observed in high energy interactions and they decay into first generation particles in short times due to their higher masses. Neutrinos of all generations do not decay and rarely interact with matter, becoming hard to detect. They oscillate between generations due to the non zero neutrino masses and neutrino mixing. Quarks, except the top quark, are bound in combinations of quarks and antiquarks called *hadrons*, with integer electrical charge. Hadrons built of three quarks are called *baryons*. Hadrons built of a quark-antiquark pair are called *mesons*.

The SM describes three of the four known forces in nature: the electromagnetic, the strong and the weak forces. At high energy scales the gravitation can be neglected, since its strength is

	Generation			Q	Spin
	I	**II**	**III**		
Quarks	up (u)	charm (c)	top (t)	$+\frac{2}{3}$	$\frac{1}{2}$
mass (MeV)	1.7-3.3	1270	172000		
	down (d)	strange (s)	bottom (b)	$-\frac{1}{3}$	$\frac{1}{2}$
	4.1-5.8	101	4190		
Leptons	electron (e)	muon (μ)	tau (τ)	-1	$\frac{1}{2}$
mass (MeV)	0.511	105.658	1776.82		
	electron neutrino (ν_e)	muon neutrino (ν_μ)	tau neutrino (ν_τ)	0	$\frac{1}{2}$
	$< 2 \times 10^{-9}$	$< 0.19 \times 10^{-3}$	< 0.0182		

Table 1.1: The fermion sector of the Standard Model. Particles are grouped into three mass generations. The electric charge Q is given in fractions of the proton charge. The masses of the quarks, except the top quark, are only estimates since no free quark can be observed. The top quark mass is measured from its decay products [13].

Interaction	Gauge boson	Q	Spin	Mass (GeV)	Particles sensitive to the interaction
Weak	$W+$	$+1$	1	80.399	Quarks and leptons
	$W-$	-1	1	80.399	
	Z^0	0	1	91.188	
Electromagnetic	γ Photon	0	1	$< 1 \times 10^{-24}$	Electrically charged leptons and quarks
Strong	g Gluon	0	1	0	Quarks

Table 1.2: The boson sector of the Standard Model. The electric charge Q is given in fractions of the proton charge. The gluon mass is a theoretical value [13].

around 43 orders of magnitude smaller than the strong interaction. Fermions interact through the exchange of bosons. The electromagnetic and the strong forces are mediated by massless bosons, photons and gluons, respectively, while the weak force is mediated by massive bosons, $W^{\pm}$ and Z^0. The last particle to complete the SM list of particles is the Higgs boson. It is the consequence of introducing a new doublet of complex scalar fields in the SM theory in order to give mass to the $W^{\pm}$ and Z^0 bosons and fermions (this will be further discussed in Section 1.2.2.3). Tables 1.1 and 1.2 list the known elementary particles and some of their properties.

1.2.2 Fundamental interactions

As mentioned before, three of the four known interactions between particles are described by the SM. They are listed in Table 1.2. Each one of them has an associated symmetry group. A fundamental property of the SM is the gauge invariance, defined as the invariance of the theory under local transformations. The SM theory is invariant under transformations of the type $SU(3)_C \times SU(2)_I \times U(1)_Y$. $U(1)_Y$ is the symmetry group of the electromagnetic interaction, $SU(2)_I$ of the weak interaction and $SU(3)_C$ of the strong interaction. C, I and Y correspond to the conserved quantum numbers for each symmetry: color charge, weak isospin and hypercharge, respectively.

1.2.2.1 The electromagnetic interaction

The electromagnetic interaction occurs between electrically charged particles. It is responsible for binding the electrons to the atomic nuclei to form atoms and then molecules. The electromagnetic force carrier is the photon γ. The photon is a massless, electrically neutral and then not self-interacting gauge boson. The electromagnetic force is described by a relativistic quantum field theory called Quantum Electrodynamics (QED). QED is based on a local symmetry (i.e. separately valid at each space-time point), called $U(1)$. As in any quantum field theory, the kinematics and the dynamics of the theory can be deduced from a lagrangian. The QED lagrangian describes the coupling between a charged fermion field ψ to the boson field A^{μ}:

$$\mathcal{L}_{\mathrm{QED}} = \bar{\psi}(i\gamma^{\mu}D_{\mu} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{1.1}$$

where γ^{μ} are the Dirac matrices. The covariant derivative D_{μ} and the field strength $F_{\mu\nu}$ are given by:

$$D_{\mu} = \partial_{\mu} - ieA_{\mu} \tag{1.2}$$
$$F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}, \tag{1.3}$$

such as the $\mathcal{L}_{\mathrm{QED}}$ is invariant under local $U(1)$ gauge symmetry ($\psi \rightarrow e^{ie\xi(x)}\psi$). The gauge invariance of the QED theory implies that the electrical charge is conserved locally. Note that the addition of a mass term for the gauge boson, of type $m^2A_{\mu}A^{\mu}$, will lead to a violation of the gauge invariance. The QED's gauge boson needs to be massless and it can be directly associated with the photon. e correspond to the elemental electrical charge and is given by:

$$e = \sqrt{4\pi\alpha_{\mathrm{QED}}}, \tag{1.4}$$

where α_{QED} is the electromagnetic coupling constant. It is a fundamental parameter of the theory and determines the strength of the electromagnetic interaction. In QED, observables are usually expressed as a function of α_{QED}. When using perturbation theory to calculate those observables, divergences appear in calculations involving Feynman diagrams with loops

including virtual particles. To avoid these divergences a method called renormalization is used. The renormalization process consists in redefining measurable observables at a given energy scale (called the normalization scale μ_0) to include the virtual particle corrections, absorbing in this way the infinities. Imposing the independence of the physical observable from μ_0 reveals that α_{QED} depends on the energy scale at which one observes the process Q^2. $\alpha_{\text{QED}}(Q^2)$ increases when the energy increases, going from $1/137$ at $Q^2 = 0$ to $1/127$ at energies corresponding to the mass of the Z boson.

1.2.2.2 The strong interaction

This interaction is responsible for holding the quarks together in hadrons and binding protons and neutrons together to form the atomic nucleus. It is described by a quantum field theory called Quantum ChromoDynamics (QCD) [14]. QCD is represented by the non-Abelian symmetry group $SU(3)$. In this representation the gluon is the gauge field, i.e. the QCD equivalent of the QED photons. Just as the electric charge in QED, QCD introduces its own charge, known as "color". Color charge comes in three varieties called red, green and blue. Antiquarks have corresponding anticolor. Quarks and antiquarks are combined in such a way that always form colorless hadrons. Leptons have no color charge. The gluon is not a charge-neutral force carrier (as its QED counterpart), it can be thought of as carrying both color charge and anticolor charge. There are eight possible different combinations of (anti)color for gluons, which form an octet in color $SU(3)$ [15]. Due to their non-Abelian nature, the gluon gauge fields exhibit self-couplings that allow for self-interactions.

The QCD lagrangian density is given by:

$$\mathcal{L}_{\text{QCD}} = \sum_q \bar{\psi}_{q,j}(i\gamma^\mu(D_\mu)_{jk} - m_q\delta_{jk})\psi_{q,k} - \frac{1}{4}G^a_{\mu\nu}G^{a\ \mu\nu}, \tag{1.5}$$

where $\psi_{q,j}$ is the quark field for flavour q and carry a color j. The covariant derivative D_μ and the gluon field strength tensor $G^a_{\mu\nu}$ are defined as:

$$D_\mu = \partial_\mu + ig_s t^a A^a_\mu \tag{1.6}$$

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_{\text{S}} f^{abc} A^b_\mu A^c_\nu, \tag{1.7}$$

where A^a_ν are the gluon fields with index a, $a = 1, ..., 8$. t^a are the matrices generators of the $SU(3)$ group, called Gell-Mann matrices. They satisfy $[t^a, t^b] = if^{abc}t^c$, where f^{abc} are the group structure constants. Finally, g_{S} is usually expressed as $g_{\text{S}} = \sqrt{4\pi\alpha_{\text{S}}}$, where α_{S} is the strong coupling constant. α_{S} has been found to have a dependence inversely proportional to the energy (after applying a renormalization process similar to the one described in the previous section). Therefore, quarks and gluons behave as quasi free particles at high energies (short distances), while at low energies (large distances) quarks are confined into hadrons. These interesting behaviors are known as asymptotic freedom and confinement, respectively. They determine the development of pp collisions and will be further discussed in Chapter 3.

1.2.2.3 The weak interaction

The weak interaction is best known for being responsible for the beta decays. The weak interaction affects all fermions, including neutrinos. It has several massive mediators, unlike the electromagnetic and strong forces, called Z^0 and $W^\pm$ bosons. It is their heaviness that accounts for the very short range of the weak interaction. The Z^0 and $W^\pm$ bosons mediate the neutral and charged weak currents, respectively. As the lifetime of a particle is proportional to the

inverse square of the coupling constant of the force which causes the decay, the lifetime of particles relying on the weak force for their decay processes is large. The weak interaction is the only interaction able to change the flavor of a quark or a lepton. In addition, it also breaks parity-symmetry since $W^{\pm}$ bosons couple only to left-handed particles, i.e. particles with spin and momentum of opposite direction, and right-handed antiparticles.

The electroweak theory The weak and the electromagnetic interactions have been successfully described as different manifestations of the same fundamental interaction by Glashow, Weinberg and Salam in the 60s. The gauge theory that describes both interactions is called unified electroweak theory and is based on the $SU(2)_I \times U(1)_Y$ symmetry group. The local gauge invariance requirement leads to the existence of four bosons: W^i_μ $(i = 1, 2, 3)$ from $SU(2)$ and B_μ for $U(1)$. The fields of the electroweak bosons (Z^μ, $(W^{\pm})^\mu$ and the photon field A^μ) are mixtures of these gauge boson fields. The lagrangian of the electroweak interaction, ruling the interaction between the gauge fields and fermions, is given by:

$$\mathcal{L}_{\mathrm{EW}} = \bar{\psi}_L(i\gamma^\mu(D_\mu))\psi_L + \bar{\psi}_R(i\gamma^\mu(D_\mu))\psi_R - \frac{1}{4}W^i_{\mu\nu}W^{i\ \mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.8}$$

where L, R refers to the left- and right-handed fermions. The gauge fields, $W^i_{\mu\nu}$ and $B_{\mu\nu}$, and the covariant derivative are given by:

$$\begin{aligned} D_\mu &= \partial_\mu + \frac{1}{2}g\tau^i_{L,R}W^i_\mu - \frac{1}{2}ig'Y_{L,R}B_\mu & (1.9)\\ W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\varepsilon_{ijk}W^j_\mu W^k_\nu & (1.10)\\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, & (1.11) \end{aligned}$$

where g and g' are the coupling constants associated to $SU(2)$ and $U(1)$, respectively. They are related to α_{QED} by $\alpha_{\mathrm{QED}} = g\sin\theta_W = g'\cos\theta_W$. θ_W is known as the weak mixing angle. The generators associated with the $SU(2)$ symmetry group are the Pauli matrices, τ_i, and the generator associated to $U(1)$ is the hypercharge, $Y = Q - I_3$, being Q the electric charge and I_3 the third component of the weak isospin. The theory as described so far predicts massless $SU(3)$ gauge fields, contradicting the experimental observations. The photon and the gluons are massless as a consequence of the exact conservation of the corresponding symmetry generators: the electric charge and the eight color charges. The fact that the weak bosons are massive indicates that the corresponding symmetries are broken. In 1964, Higgs, Brout and Englert proposed that the breaking of the electroweak gauge symmetry is induced by the Brout-Englert-Higgs mechanism, which predicts the existence of a spin 0 particle, known as the Higgs boson, not yet experimentally observed. The Brout-Englert-Higgs mechanism consists in introducing an additional complex scalar doublet, $\Phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix}$, where:

$$\begin{aligned} \phi^+ &\equiv (\phi_1 + i\phi_2)/\sqrt{2} & (1.12)\\ \phi^0 &\equiv (\phi_3 + i\phi_4)/\sqrt{2}. & (1.13) \end{aligned}$$

The Higgs lagrangian is given by:

$$\begin{aligned} \mathcal{L}_{\mathrm{Higgs}} &= (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi^\dagger\Phi) & (1.14)\\ V(\Phi^\dagger\Phi) &= \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2, & (1.15) \end{aligned}$$

where D_μ is given by Eq. 1.9. For $\mu^2 < 0$ and $\lambda > 0$, V has its minimum at $\Phi^\dagger\Phi = -\frac{\mu^2}{2\lambda}$. Spontaneous symmetry breaking is induced if the minimum of V is obtained for non-vanishing Φ values. Expanding Φ around a particular minimum, say:

$$\phi_1 = \phi_2 = \phi_4 = 0 \tag{1.16}$$

$$\phi_3^2 = -\frac{\mu^2}{\lambda} \equiv \upsilon^2, \tag{1.17}$$

we obtain:

$$\Phi = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ \upsilon + H \end{pmatrix}. \tag{1.18}$$

Of the four scalar fields only the Higgs field H remains. The other three scalar fields become the longitudinal modes of the $W^\pm$ and Z which acquire a mass. The value of gauge bosons masses can be found by their coupling to the Higgs field. By replacing Φ by Eq. 1.18 into $\mathcal{L}_{\text{Higgs}}$:

$$m_W = \frac{1}{2}\upsilon g \tag{1.19}$$

$$m_Z = \frac{1}{2}\upsilon\sqrt{g^2 + g'^2} \tag{1.20}$$

$$m_\gamma = 0. \tag{1.21}$$

Similar to the generation of the gauge boson masses, the fermion masses can be introduced. The coupling between the Higgs field and massless quark and lepton fields are described by the Yukawa interactions. Fermions acquire a mass proportional to υ. But the Higgs boson has not been experimentally observed yet. A huge experimental effort is underway at LHC to reveal the Higgs sector. Its mass, $m_H = \sqrt{2\lambda}\upsilon$, is not predicted by the SM since λ is a free parameter. The searches at LEP, Tevatron, ATLAS and CMS have set experimental limits on its mass. Apart from the experimental constraints, there are theoretical bounds on the value of the Higgs mass. Both, theoretical and experimental constraints in the Higgs boson mass will be briefly discussed in Section 1.2.3.

The last (important for the studies presented here) missing point corresponds to the discussion of the fermion flavor changes. Weak charged currents are the only interaction in the SM that changes the flavor of the fermions: for example, by emission of a W boson an up-quark is turned into a down-quark, or a ν_e neutrino is turned into an e^-. The mass eigenstates of fermions are not identical to the weak eigenstates. The transformation between them is described by a 3×3 unitary matrix: the Cabibbo-Kobayashi-Maskawa (CKM) matrix describes the mixing of the quark eigenstates, while the Maki-Nakagawa-Sakata (MNS) matrix describes the mixing for leptons. The CKM matrix is given by (as it will be referred to in the future) [13]:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.97428 & 0.2253 & 0.00347 \\ 0.2252 & 0.97345 & 0.0410 \\ 0.00862 & 0.0403 & 0.999152 \end{pmatrix}. \tag{1.22}$$

The probability for a quark of flavor i to be transformed to a quark of flavor j, emitting a W boson is proportional to $|V_{ij}|^2$. The CKM and MNS matrix elements are free parameters of the SM and need to be determined experimentally.

1.2.3 Limitations of the Standard Model

Despite its predictive and descriptive power, the SM has several limitations of experimental and theoretical nature. This section contains an overview of the most important known problems of the SM.

Theoretical limitations

- The SM theoretical framework does not describe the gravitational force, which becomes important at very high energies, i.e. small length scales quantified by the Planck scale $M_{\text{Planck}} \approx 10^{19}$ GeV. The Planck scale is seen by many as the limit of the SM validity. So far, no consistent quantum field gravity theory has been constructed.

- In the most basic formulation of the SM, particles are massless. In order to give mass to particles the spontaneous symmetry breaking mechanism needs to be included in an *ad hoc* and artificial way. The SM does not explain the origin of this mechanism, except for its need in order to match the experimental observations.

- Another indication of the incompleteness of the SM is the *hierarchy problem*. Due to divergent loop corrections to the Higgs mass, a renormalization process needs to be applied to calculate the Higgs mass. At first order, the Higgs mass can be written as:

$$m_H^2 = (m_H^2)_0 - \frac{\lambda_f^2 \Lambda^2}{8\pi^2}, \tag{1.23}$$

  where the first term corresponds to the bare Higgs mass squared and the second term corresponds to one-loop quantum corrections at first order involving a fermion. λ_f corresponds to the coupling constant of the Higgs with the fermion, known as the Yukawa coupling. The size of the correction depends on Λ, the scale of the process. If we consider that the SM is valid up to the Planck scale, $\Lambda = M_{\text{Planck}}$, and $m_H \approx 100$ GeV is required, then there has to be an "unnatural" fine-tuning to balance the correction term with respect to the first term. This procedure cast doubt upon the robustness and universality of the theory. The hierarchy problem is a consequence of the difference between the strengths of the electroweak and the gravitational forces.

- The SM does not explain the number of fermion families observed so far, nor its mass hierarchy.

Experimental limitations

- Experimental astrophysical observations of the rotation curves of galaxies indicate that visible matter constitutes around 17% of the mass in the universe. The rest of the mass is believed to consist of massive, weakly interactive and stable particles called "dark matter" particles. Dark matter particles combined with visible particles make up the 30% of the energy in the universe [16]. The rest is presumably composed by "dark energy", which is used to explain the acceleration of the universe [17]. Neither dark matter nor dark energy nature are included in the Standard Model theory.

- In the SM, the neutrino is massless. However, experimental neutrino's oscillations measurements indicate that neutrinos have a finite but small mass [18].

- The non-observation of the Higgs boson can also be regarded as a weakness of the SM. Experimental results from LEP set a lower mass limit for the Higgs mass to 114.4 GeV [19].

At the beginning of 2012, the combined results from DØ and CDF at Tevatron excluded the 147-179 GeV and 100-106 GeV ranges at a 95% of confidence level and found an excess of data with a local significance of 2.2σ in the region of 115 to 135 GeV [20]. Recent results from the ATLAS experiment excludes at 95% CL the following ranges: 110-117.5, 118.5-122.5, 129-539 GeV and found an excess with a local significance of 2.5σ around 126 GeV [21]. While the CMS experiment results exclude the range between 127.5 GeV and 600 GeV at 95% CL and the largest observed excess is at 125 GeV with a local significance of 2.8σ [22]. Eventhough the SM does not predict the Higgs mass, it is possible to constrain it using the measured masses of the weak bosons and the top quark. Electroweak precision measurements like the weak bosons and the top quark masses constrain the loop radiative corrections caused by the Higgs boson and therefore can be used to put indirect limits on the Higgs mass. The latest results indicate that when including the results from direct Higgs searches from 2010 at LHC and Tevatron, the upper limit on the Higgs mass is 143 GeV at 95% CL. This limit is alleviated to 169 GeV when not including the direct Higgs searches [9]. Figure 1.1 shows the observed and expected limits at 95% CL, as functions of the Higgs boson mass from Tevatron studies. The regions excluded for LEP and LHC are also shown for reference (they do not correspond to the latest results from LHC, but to the results known at February 2012)[1].

1.3 Beyond the Standard Model

The SM has proved to effectively describe the interactions of known matter at the electroweak energy scale. However, it presents some limitations that have inspired many physicists to develop several theories to correct the weaknesses of the SM. So far, none of them have managed to cover and/or solve all the limitations of the SM. Those theories are known as Beyond-the-SM (BSM) theories and their verification/refutation is also one of the goals of the LHC physics program. The general believe is that the SM is the low energy approximation of another theory that unifies all the four interactions in nature. So far, no unification theory has been found. But many theories have been proposed to try to account for one or more of the SM's limitations. In this section, a brief overview of the motivation, features and phenomenology of some BSM theories are presented.

1.3.1 Supersymmetry

Supersymmetry (SUSY) [25, 26] offers a solution to the hierarchy problem through the stabilization of the mass scale in the Higgs potential. SUSY is a symmetry that allows transformation between fermions and bosons, i.e. bosons and fermions form supersymmetric multiplets. In the simplest supersymmetry model known as Minimal Supersymmetric Standard Model (MSSM) each SM fermion has a supersymmetric partner boson (a particle with the same mass and quantum numbers but the spin, which differs by 1/2) is added and vice-versa. MSSM has around 124

[1] On July 4^{th} 2012 ATLAS and CMS presented their latest results on the search for the SM Higgs boson using the data recorded up to June 2012. ATLAS results exclude the SM Higgs boson at 95% CL in the mass range 111-559 GeV, except for the narrow region 122-131 GeV, where an excess with a local significance of 5.9σ is observed mainly in the $H \to ZZ^{(*)} \to 4l$, $H \to \gamma\gamma$ and $H \to WW^{(*)} \to l\nu l\nu$ channels. This results in a clear evidence for the production of a neutral boson with a measured mass of $126.0 \pm 0.4(\text{stat}) \pm 0.4(\text{sys})$ GeV [23]. CMS obtained similar results. CMS data exclude the existence of a SM Higgs boson in the ranges 110-122.5 and 127-600 GeV at 95% CL and shows an excess around 125 GeV with a local significance of 5.0σ and a measured mass of $125.3 \pm 0.4(\text{stat}) \pm 0.5(\text{syst})$ GeV [24]. Both results are compatible with the expectations of the SM Higgs boson. However, more data is be needed to check more precisely if this new neutral boson has all the properties of the SM Higgs boson.

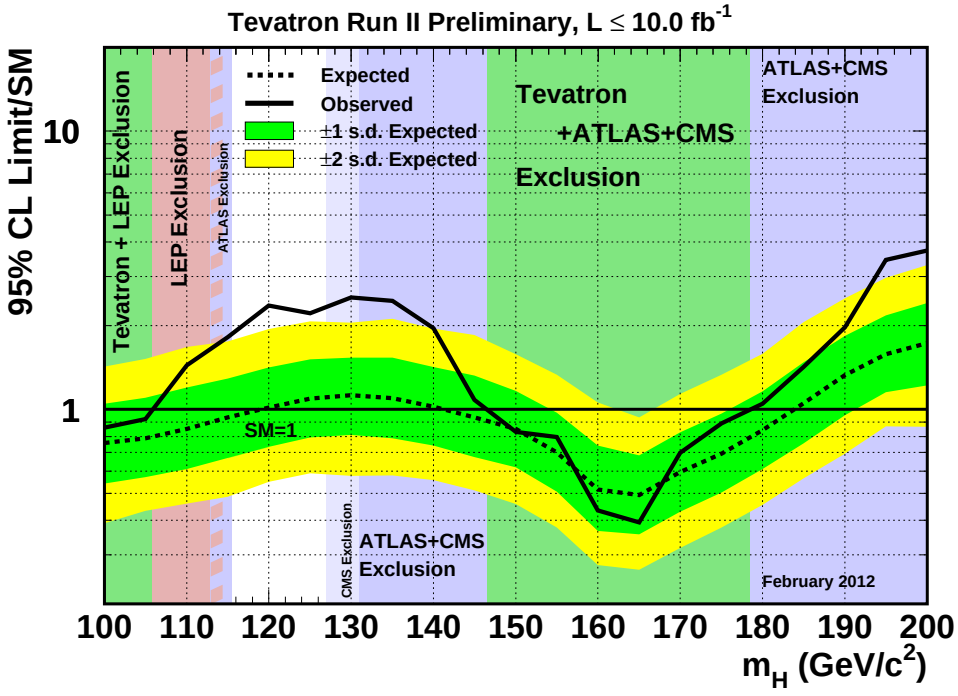


Figure 1.1: Observed and expected limits at 95% CL, as functions of the Higgs boson mass for the combined CDF and DØ analyses. The bands indicate the 68% and 95% probability regions where the limits can fluctuate, in the absence of signal. The regions excluded by other experiences are also shown. The ATLAS and CMS exclusions in the plots are not updated [20].

free parameters instead of the 19 present in the SM.

There are two main theoretical motivations for SUSY. The first one regards the hierarchy problem. Supersymmetry has been proved to be free from quadratic divergences. The second term in Eq. 1.23 coming from SM bosonic and leptonic loops is canceled by the supersymmetric partners loops. In this way, the scale of validity of the theory is extended without introducing the hierarchy problem. However, supersymmetric particles have not been observed yet, implying that they do not have the same mass as their corresponding SM partners. If supersymmetry exists, then it has to be a broken symmetry. After breaking supersymmetry the quadratic divergences cancel but the radiative corrections in Eq. 1.23 are not completely canceled. However, they remain negligible if the introduced supersymmetric particles are not much heavier than the TeV scale, raising the possibility of testing the MSSM model at the LHC. The second one is the fact that SUSY theory allows the approximate unification of the coupling constants of the electromagnetic, weak and strong interactions at high energies. Using supersymmetry the three coupling constants converge at an energy scale of around 10^{16} GeV.

Other interesting features of SUSY are of experimental nature. In supersymmetry the spontaneous electroweak symmetry breaking appears in a natural way, driven by quantum radiative effects (after imposing the soft supersymmetry breaking). It is obtained by the Brout-Englert-Higgs mechanism. Supersymmetry requires two Higgs doublets, instead of the single Higgs doublet of the SM. From the eight degrees of freedom of the two complex doublets three are eaten in the Brout-Englert-Higgs mechanism and five correspond to physical particles: two real CP-even scalars, one real CP-odd scalar and one complex scalar. The introduction of the supersymmetry breaking is however still arbitrary. Supersymmetry also provides a dark matter

particle candidate. Some SUSY models introduce a new quantum number, called $R-$parity and denoted as R, which is conserved. R is -1 for SUSY particles and 1 for SM particles. If R is conserved any SUSY particle has to decay to at least one other SUSY particle and they should be produced in pairs. The lightest SUSY particle should be stable, with a mass of the TeV order and could be a candidate to explain dark matter, since it interacts weakly with the matter.

The evidence for SUSY will arrive basically from the discovery of the supersymmetric particles partners of the SM. They usually produce long decay cascades of particles which include leptons and jets, with very high multiplicities. In the case of R conserving SUSY models the usual signature of final states is the presence of a high missing transverse energy and momentum, as the lightest supersymmetric particle will escape the detector, leaving an unbalanced momentum in the transverse plane.

1.3.2 Technicolor and topcolor

Technicolor [27, 26] models provide a different solution to the hierarchy problem by removing the fundamental scalar particles from the theory. So, instead of introducing Higgs bosons to break the electroweak symmetry, technicolor models introduce a new force, similar to the strong force, and additional massless fermions, called technifermions, that feel this new force. The new force, known as technicolor, becomes strong at a scale of around 500 GeV, leading to the formation of technifermions condensates. Because the left-handed technifermions carry electroweak quantum numbers, but the right-handed do not, the formation of these condensates breaks electroweak symmetry. Also, in these models the chiral symmetry of the massless technifermions is broken when the technifermions condensates form. Therefore, the technifermions acquire a dynamical mass and three composite massless Goldstone bosons are generated and give mass to the W and Z bosons. So far, the model explains the generation of the electroweak gauge bosons masses, but not of the quarks and leptons masses. In order to generate massive SM fermions new interactions need to be added to the model. They are known as Extended Technicolor (ETC) forces and couple SM fermions with technifermions. The scale of ETC symmetry breaking is estimated to be of the order of 100 TeV, using experimental measurements, which means that ETC can not generate fermions with a high mass as the top quark. The experimental signature of the ETC model is the presence of new strong-interacting scalar resonances of technifermion-antitechnifermions pairs.

One extra extension of technicolor models is their combination with the Topcolor model [28]. The Topcolor model explains the large top quark mass through the introduction of a new strong gauge interaction called topcolor. Topcolor becomes strong near 1 TeV. In this model the QCD gauge group, $SU(3)_C$, is included into a larger symmetry $SU(3)_1 \times SU(3)_2$. $SU(3)_2$ ($SU(3)_1$) couples to the third (second and first) generation. The breaking $SU(3)_1 \times SU(3)_2 \rightarrow SU(3)_C$ produces massive color octets gauge bosons, called colorons, which couple preferentially to the third generation quarks and weakly with the rest of the families. It enhances the formation of $t\bar{t}$ and $b\bar{b}$ condensates. By itself, this scheme predicts large fermions masses of about 600 GeV. The observed top quark mass is assured when topcolor is combined with an additional strong dynamics technicolor.

This combined theory is known as Topcolor Assisted Technicolor [29, 30]. In this framework the electroweak interaction is broken by ETC. The light quarks and leptons masses are generated by ETC. Finally the top quark high mass is explained by the fact that it is a combination of a $t\bar{t}$ condensate generated by the topcolor force, together with a small fundamental component generated by ETC. In order to keep the $b-$quark light in comparison to the top quark, the production of $t\bar{t}$ condensate needs to be enhanced against the $b\bar{b}$ condensate production. There are several ways to do so, but the simplest one is to include the $U(1)$ symmetry into a largest one as $SU(1)_1 \times SU(1)_2$. The breaking $SU(1)_1 \times SU(1)_2 \rightarrow U(1)$ generates a massive neutral gauge

boson, Z', which couples preferentially to the third generation quarks. Z' is a colour singlet resonance with spin 1.

1.3.3 Theory of extra dimensions

The introduction of extra dimensions was first developed by Kaluza [31] and Klein [32], to unify electromagnetism and relativistic gravity. Extra dimensions appear in several BSM theories. For example, in string theory they are used to achieved the quantization of the gravitational interaction. Later on, Randall and Sundrum [33] went further proposing a model with a single extra warped (curled up with a finite radius) extra dimension to explain the hierarchy problem. In the original model gravity propagates in the added dimension, while the SM fields are contained in our normal four-dimensional spacetime. In an extension to the model, all the SM fields are free to propagate in the extra dimension too. This extension has desirable features as the suppression of FCNC, a natural explanation of the fermion masses hierarchy and gauge coupling unification.

The hierarchy problem is solved geometrically through an exponential warp factor. The metric of the warped space can be written as [34, 35, 36]:

$$ds^2 = e^{-k|y|}\eta_{\mu\nu}dx^\mu dx^\nu - dy^2, \tag{1.24}$$

where y represents the extra dimension, k determines the curvature of the space and $e^{-k|y|}$ is the warp factor. The space has boundaries at $y = 0$ and $y = \pi r_c$, r_c is the radius of the warped dimension. The boundary at $y = 0$ is called the *Planck brane*, while the one at $y = \pi r_c$, r_c is called the *TeV brane*. The particles of the SM are on the TeV brane. And the gravitation resides on the Planck brane. Choosing $kr_c \approx 11$ then the ratio of the Planck to TeV scales is:

$$\frac{e^{-k|y(Planck)|}}{e^{-k|y(TeV)|}} = e^{-k\pi r_c} \approx 10^{-15}. \tag{1.25}$$

Said on other way: the fundamental Planck scale is actually small in the Planck brane ($\approx 10^{19}$) and so the gravity is actually strong, while in the TeV brane it is of the order of TeV and the gravity appears as a weak interaction, i.e solving the hierarchy problem. When a SM particle is excited into the extra dimension, it acquires an effective mass in the 4-dimensional spacetime, generating the so called Kaluza-Klein excitations. Kaluza-Klein (KK) excited states of the gluons and electroweak gauge bosons, as well as graviton are predicted by this model. Of particular interest are the KK-gluons, since they have the higher production rate at pp colliders and are likely to be detected at present high energy experiments. The KK-gluons are expected to decay dominantly into $t\bar{t}$ final states. Their mass is estimated to be of the TeV order, their decays give energetic top quarks final states. The coupling between the SM fermions and the KK-gauge bosons, determine the phenomenology of the model. Eventhough this model is able to explain some of the SM limitations, some fine-tuning needs to be applied in order to match some electroweak precision measurements.

Another kind of extra dimension models is the Arkani-Dimopoulos-Dvali (ADD) model [37]. It requires that the SM fields are contained in our normal four-dimensional space, while gravity propagates through several spatial extra large (larger than the Planck's scale) dimensions. In this model, the fundamental scale is much lower than the Planck's scale. This occurs because the power law of gravity changes. With the correct selection of the number and size of extra large dimensions the Planck's scale can be reduced to the TeV order (close to the electroweak scale order). Then, the hierarchy problem is replaced by the problem of choosing the right number and size of extra large dimensions. One experimental signature in the ADD model is the creation of microscopic black holes. The existence of large extra dimensions can increase the value of G, the gravitational constant. So, if there are a few number of large extra dimensions, the collisions

at the LHC should produce microscopic black holes, through the squeezing of a pair of partons below their combined Schwarzschild radius. The production of microscopic black holes have been hardly constrained by CMS [38] and ATLAS [39] experiments studies.

1.3.4 String theory

So far, all the attempts to incorporate gravitation in the SM have failed because the general relativity and the quantum field theory seems to be mathematically incompatible, since general relativity is not renormalizable (due to the fact that SM particles are 0-dimensional objects). The string theory [40] attempts to reconciliate both theories by assuming that fermions and bosons are 1-dimensional objects, called strings. Strings can oscillate giving in this way their flavor, charge, mass and spin to the observed particles. The quantification of the theory leads to the existence of a spin 2 boson, which is identified as a graviton. Another particle with imaginary mass, called the tachyon, appears too in the modelization. To avoid this mathematical inconsistency, the supersymmetry is included in the string theory. This new theory is called superstring theory. String theory predicts extra dimensions: 26 for the original string theory and 10 for the superstring. One explanation offered is that 6 of these extra dimensions are so small that so far remain undetectable by experiments.

The characteristic scale of the string theory is expected to be close to the Planck's scale. At low energies below the Planck's scale, experiments can not resolve short distance of the order of the Planck's length ($\approx 1.6 \times 10^{-33}$ cm) and strings can be seen as 0-dimensional objects. This could explain the success of using quantum field theories so far. String theory's techniques have been used to describe qualitative attributes of the quark-gluon plasma with good results. It also describes black hole's thermodynamics with a good accuracy. However, the definitive test of the string theory is a difficult task. First, since the Planck's length is very small, the structure of string would be hardly tested in current high energy experiments. Second, there are a big number of string theories with different configurations. The most accessible experimental implication (but not definitive) of superstring theory is supersymmetry. String theory is still a work in progress. There are some things not yet understood, for example the emerging of the SM at low energies is not included yet. However, this theory shows great promise as a unified quantum theory of all fundamental forces known so far.

1.4 The top quark

The top quark is the heaviest known elementary particle. Its mass is five orders of magnitude larger than the mass of the first generation quarks and is close to the electroweak symmetry breaking scale. The study of its properties will allow to test the predictions of the SM. On the other hand, its high mass could hint an intimately connection between the top and BSM physics. Given its important role in SM and BSM theories, some of its particular properties are reviewed here.

The top quark was the last quark to be discovered. It was first directly observed in $p\bar{p}$ collisions at DØ and CDF experiments at the Tevatron accelerator at Fermilab in 1995 [41, 42]. However, its existence was already predicted by the SM before its direct observation. The $b-$quark was discovered in 1977. It decays through charged currents, which is only allowed for $SU(2)_L$ doublets, implying that the weak isospin partner of the $b-$quark should exist. In addition, the existence of the $b-$quark isospin partner would avoid the existence of divergences arising sometimes in triangle Feynman diagrams, also known as triangle anomalies. The top mass was also predicted with a high precision. The mass of the W boson m_W can be expressed in terms of the electromagnetic coupling constant α_{EM}, the Fermi constant G_F and the mass

of the Z boson m_Z. Radiative corrections to m_W are dominated by virtual top quark loops. The top quark high mass makes this contribution much bigger than the ones from bosons and other fermions. Physical quantities can be calculated using perturbative theory to a given order of precision. The first order of calculation is known as leading order (LO), which involves just basic diagrams without loops or extra vertices. When including radiative corrections, i.e an additional level of complexity of loops and vertices, the calculations are done to higher orders: next to leading order (NLO) or next to next to leading order (NNLO). With each extra order of corrections the calculation becomes more precise but also more difficult and longer. Thus, m_W was used to calculate the predicted top mass. This theoretical prediction agrees with the best experimental measurement of the top mass obtained by Tevatron within a good precision:

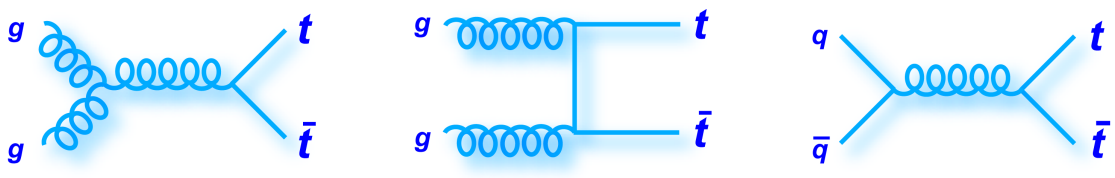
$$m_t^{\text{Tevatron}} = 173.2 \pm 0.9 \text{ GeV [43]}$$
$$m_t^{\text{predicted}} = 173.4 \pm 1.1 \text{ GeV [9]}.$$

A complete summary of the latest results from top properties studies at Tevatron and the LHC experiments were presented in the Moriond conference 2012 and can be found in Ref. [44]. The production of the top quark in the SM framework as well as its decay modes will be briefly described in the following sections.

1.4.1 Top quark production

1.4.1.1 Top pair production

The dominating production mechanism for top quarks at hadron colliders is via the strong interaction. As the strong interaction conserves the flavour, the top quark is produced in this case by pairs: 82% through gluon fusion and 18% through quarks annihilation at the LHC at $\sqrt{s} = 7$ TeV. At the LHC at $\sqrt{s} = 7$ TeV, the NNLO approximate cross section of top pair production is around 163 pb [45]. Figure 1.2 shows the Feynman diagrams for the top pair production at LO.



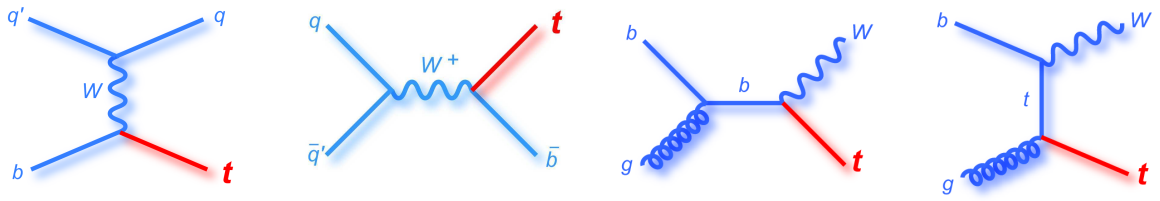
(a) (b) (c)

Figure 1.2: Feynman diagrams of top pair quark production by (a,b) gluon-gluon fusion and (c) quark-antiquark annihilation.

1.4.1.2 Single top production

At hadron colliders, the secondary mechanism of top quark production is the single top production, which is mediated by the electroweak interaction. There are three production channels which are shown in Figure 1.3 at the leading order. The first diagram, known as $t-$channel, is the dominant process, with a NNLO approximate cross section of around 64.57 pb at the LHC at 7 TeV [46]. The second diagram corresponds to the so called $s-$channel, while the last one is

known as the $Wt-$channel. Their calculated NNLO cross section at the LHC at 7 TeV are 4.63 pb [47] and 15.74 pb [48], respectively.



(a) (b) (c)

Figure 1.3: Feynman diagrams of single top quark production in the (a) $t-$channel, (b) $s-$channel and (c) $Wt-$channel.

1.4.2 Top quark decay

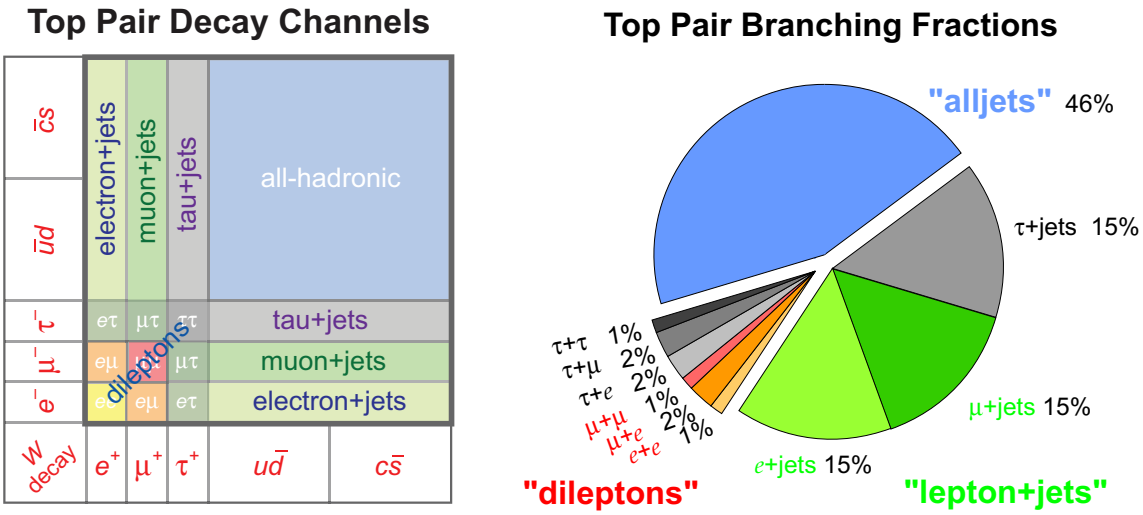
The estimated top quark lifetime is around 5×10^{-25} s. On the other hand, the characteristic hadron formation time is $\approx 3 \times 10^{-24}$ s. This means, that the top quark decays before hadronizing and has to be detected through its decay products. In the SM, the only possible top decays are: $t \to bW^+$, $t \to sW^+$ and $t \to dW^+$. The probability of these decays to occur is proportional to the corresponding elements of the CKM matrix, $|V_{tq}|^2$ with $q = b, s, d$, respectively. Therefore, the top quark decays mainly into a W boson and a b quark with a branching ratio of $BR(t \to bW^+) = 0.99^{+0.09}_{-0.08}$ [13]. The top decay final states are determined by the decay of the W, which decays approximately 33% of the times into a charged lepton-neutrino (leptonic decay) pair and 67% into a quark-antiquark pair (hadronic decay). The exact branching fractions for the different W decay channels are summarized in Table 1.3. If the lepton is a τ, it will decay subsequently into an electron/muon-neutrino pair or hadronically.

Channel	Branching fraction
$e + \nu_e$	$(10.75 \pm 0.13)\%$
$\mu + \nu_\mu$	$(10.57 \pm 0.15)\%$
$\tau + \nu_\tau$	$(11.25 \pm 0.20)\%$
$q\bar{q}$	$(67.60 \pm 0.27)\%$

Table 1.3: Branching fractions for the different W boson channel decays [13].

This results in three types of decays for top pairs, namely: *hadronic*, *lepton plus jets* (also known as *semileptonic*) and *dileptonic*. The corresponding branching ratios follow from the individual branching fractions of the W boson decay modes and are summarized in Figure 1.4. Example of each type of decay are schematically illustrated in Figure 1.5. In Chapter 2 the LHC machine and the ATLAS detector are described. This will help us to understand the detector response to the different final states of the top quarks events, to identify their detector signature and the experimental advantages and inconvenients of each top pair decay channel.

The LHC is the second particle collider observing the top quark. Only during 2011 the LHC delivered more than 10^6 top quarks. At the LHC the study of the top quark is important for two main reasons. The first one is to test the SM predictions. As discussed in Section 1.2.3 a precise top quark mass measurement can be used to constraint the SM Higgs boson mass, the



(a) (b)

Figure 1.4: (a) Top pair channel decays and (b) their approximate branching ratios.

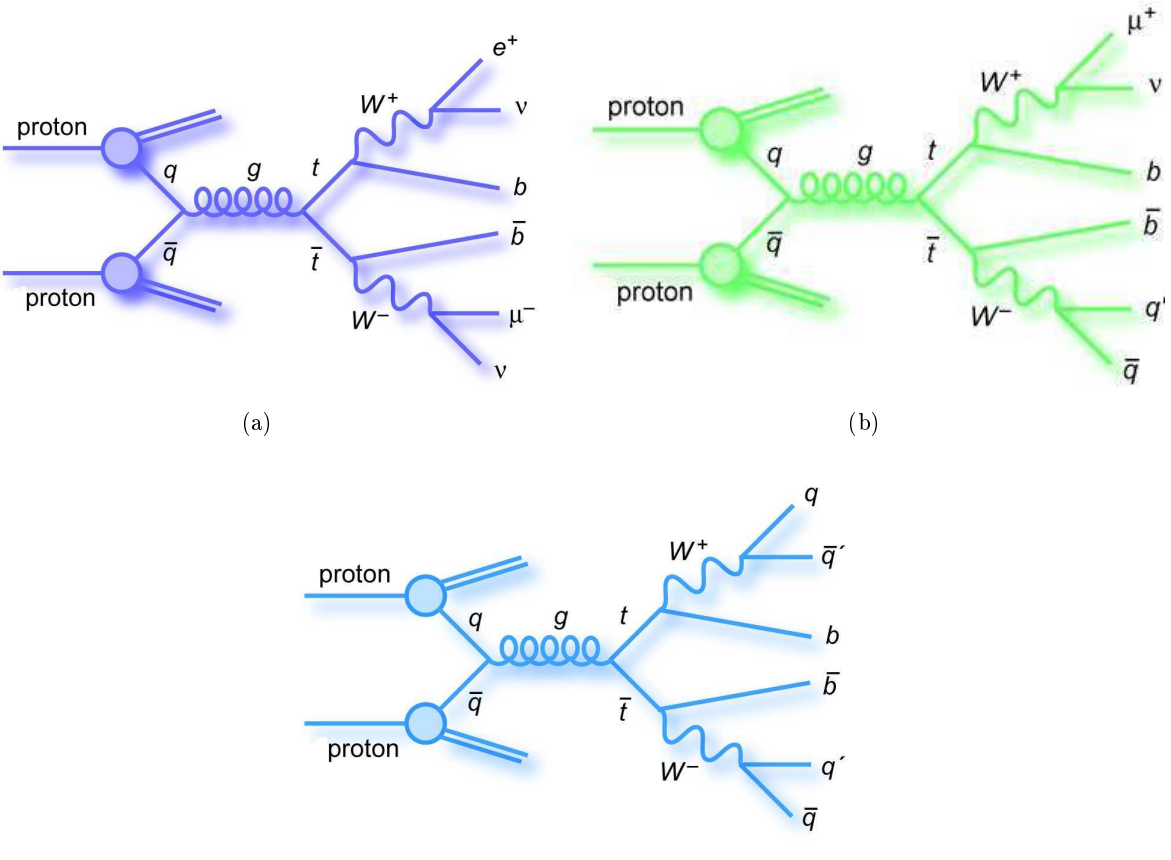


Figure 1.5: Feynman diagrams examples of top quark pair decay in the (a) leptonic, (b) lepton plus jets and (c) all hadronic channel.

only missing particle in the SM framework. To verify the SM it is also important to measure the properties of the top quark, others than the mass, as its cross section, charge and spin. Any

deviation of these properties with respect to the SM predictions would imply the interference of new physic phenomena.

Leading to the second reason: the top quark plays an important role in BSM theories. In the last year the search for new physics in the top sector has been encouraged by the recent top forward-backward asymmetry measurement at Tevatron, which found for $t\bar{t}$ masses above 450 GeV a positive asymmetry that disagrees with the SM at 3.4σ [49]. Many models (see Section 1.3 for reference) predict the existence of new particles that couple preferentially to the top quark. $t\bar{t}$ production seems to be a good and natural place to look at. If there is a new particle decaying into top pairs, one of its most important background will be the SM top production described above, which will constrain the possible extraction of a new physics signal. The results of the search for $t\bar{t}$ resonances are described in Chapter 6.

Chapter 2

The ATLAS detector

Contents

2.1 The large hadron collider

The Large Hadron Collider (LHC) [50] is the world's largest and highest energy particle accelerator, a genuine work of art of engineering. This circular particle accelerator is situated at CERN, at the Franco-Swiss border near Geneva in Switzerland, in a tunnel of 27 kilometers in circumference that can reach a depth of 175 meters. The LHC was designed to accelerate protons up to energies of 7 TeV and produce proton-proton (pp) collisions at a high rate [1]. The counter-rotating proton beams cross each other in four points along the tunnel, where the particle detectors ALICE [51], ATLAS [52], CMS [53] and LHCb [54] are located. Figure 2.1 locates the detectors along the LHC ring as well as the pre-accelerators: the linear accelerator (LINAC), the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS).

2.1.1 Design LHC running conditions

In this section the full LHC running chain at the design conditions is described. Protons are created from hydrogen atoms ionized by an electric field. Then they are accelerated by a serie of accelerators that progressively increase their energy, starting by the linear accelerator LINAC 2 which raises protons' energy up to 50 MeV. The protons are then accelerated in three steps by the circular pre-accelerators to 1 GeV (PSB), 26 GeV (PS) and 450 GeV (SPS) respectively. 450 GeV is the minimum energy at which LHC can maintain a stable beam. An arrangement

[1] Occasionally beams of heavy ions are accelerated and collided instead.

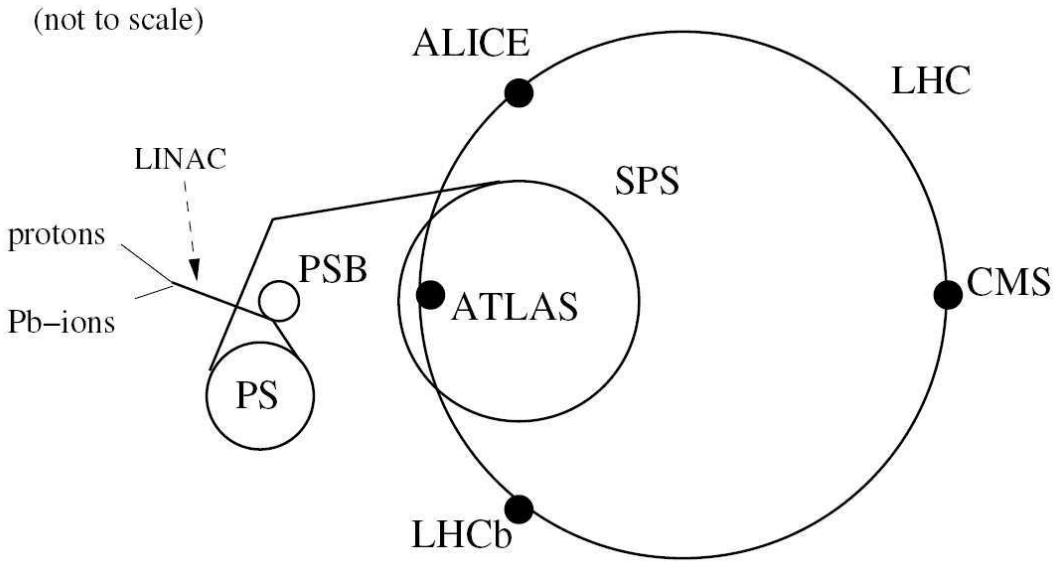


Figure 2.1: Schematic view of the particle accelerators and detectors at CERN.

of protons into bunches of 1.15×10^{11} protons is established at the nominal conditions. These bunches are organized in several "bunch trains". The spacing between bunches within a bunch train is approximately 25 ns at the design conditions, while bunch trains are further apart. The minimum distance between bunches is about 7 m, which for a 27 Km tunnel gives approximately 3550 bunches. However the effective number of bunches is 2808 in order to leave room for beam injection and abort procedure.

After protons are organized into bunches the SPS injects them into the LHC, both in clockwise and counter-clockwise directions. Then the LHC accelerates the protons to energies up to 7 TeV (the whole protons acceleration process takes around 25 minutes). These bunches circulate in separate vacuum tubes in opposite directions. Those tubes are surrounded by several thousands of superconducting magnets, which accomplish the bending and focusing of the beams. As the radius of the accelerator is fixed by the existing tunnel, the energy of the proton beams is constrained by the strength of the bending magnets. The bending is achieved by 1232 dipole magnets, which are cooled to 1.9 K by liquid helium and provide a field strength of 8.33 T at most.

Once beams are accelerated to the desired energy, stable beams are declared and the LHC experiments can start taking data usable for analysis. The intensity of beams decreases with time due to a loss of protons through collisions or through other effects influencing the trajectories of individual protons in the ring. It results in an expected beam lifetime of approximately 10 hours at the design conditions. When the intensity of the beam is too low, it is "dumped" or directed out of the accelerator into a large metal block where it is absorbed. The dipole magnets are then ramped down to 0.54 T and they stay at flat bottom for some 20-40 minutes. Meanwhile beam injection is repeated before the magnets are ramped up again to 8.3 T for another cycle of high energy collisions. This beam cycle is called a "fill". The machine is designed to withstand some 20 000 such cycles in a 20 years' lifetime [55].

2.1.2 Early LHC operation

The LHC produced its first pp collisions on November 23, 2009 at the injection energy of 450 GeV. Less than a week later, it became the most powerful collider in the world, reaching beam energies of 1.18 TeV. The energy was gradually ramped up to reach a center of mass energy of 2.36 TeV on December 8, 2009. On March 19, 2010 the LHC broke a new record by raising the beam energy to 3.5 TeV, and the first pp collisions at this energy were recorded on March 30, 2010. The beam energy for the whole 2011 year was 3.5 TeV per beam, while in 2012 the beam energy is 4 TeV. At the beginning of 2013 the LHC will go into a long shutdown to prepare for higher energy collisions starting in 2014. Results presented in Chapter 5 use pp collisions data at a center of mass energy of 7 TeV collected in 2010 while Chapters 6 and 7 use data collected in 2011. Several distinct periods of machine configuration and detector operation were present during the 2010 and 2011 data taking. The spacing between the bunches was no less than 150 ns in 2010 and it was reduced to 50 ns in 2011.

2.1.3 LHC luminosity and pile-up

An important characteristic of an accelerator machine is the instantaneous luminosity L. It allows determining the rate of pp interactions and thereby the rate of interesting events of a given process that occur in the center of the ATLAS detector:

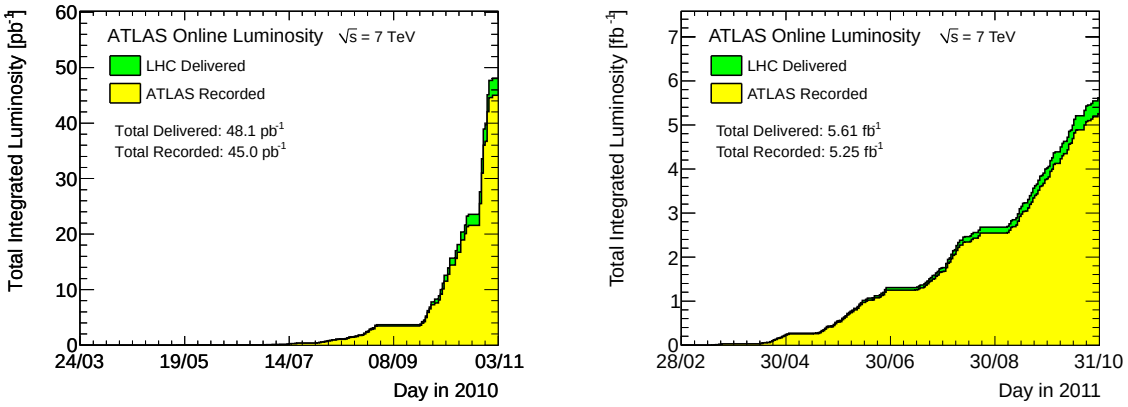
$$\frac{dN_{\text{events}}}{dt} = L\sigma_{\text{events}}, \tag{2.1}$$

where σ_{events} is the cross section and N_{events} is the number of those events at a given center of mass energy $\sqrt{s}$. The maximization of the delivered luminosity is important to observe rare physics processes (with low cross section). This can be done by squeezing the beams and reducing their transverse size or by increasing the number of bunches or the number of protons in each bunch. In nominal conditions the instantaneous luminosity delivered by the LHC will be $L = 10^{34}$ cm^{-2} s^{-1}. The instantaneous luminosity is measured using ATLAS subdetectors built for this purpose [56, 57] and described in Section 2.2.5. What is typically quoted is not the instantaneous luminosity but the integrated luminosity, integrated in a given time period. It is estimated to be 100 fb^{-1} per year when running at design instantaneous luminosity. As the machine has not yet reached the nominal conditions the integrated luminosity recorded in the past two years (which correspond to the data used in this thesis) is lower: 48.1 pb^{-1} in 2010 and 5.61 fb^{-1} in 2011 (see Figure 2.2). The systematic error on the integrated luminosity for the full 2010 (2011) dataset is 3.4 (3.9)%.

The probability of having multiple pp interactions in each event increases proportionally with the instantaneous luminosity. This phenomena is known as pile-up. There are two kinds of pile-up events: in-time pile-up corresponds to extra pp collisions within the same bunch crossing, while out-of-time pile-up corresponds to additional proton-proton interactions ocurred in a previous bunch crossing. The out-of-time pile-up is the result of long electronic integration times within the detector and becomes important when the bunch crossing spacing decreases.

The experimental observable used as an estimator of the in-time pile-up is the number of reconstructed primary vertices N_{PV}. To estimate the out-of-time pile-up the average number of pp collisions per bunch crossing at the time of the recorded event $< \mu >$ is used. $< \mu >$ is calculated using the average luminosity L over a large time period Δt ($\Delta t \gg 600$ ns), the total inelastic pp cross section σ_{inel}, the number of colliding bunches in LHC N_{bunch} and the LHC revolution frequency f_{LHC} [56]:

$$< \mu >= \frac{L \times \sigma_{\text{inel}}}{N_{\text{bunch}} \times f_{\text{LHC}}}. \tag{2.2}$$



(a) (b)

Figure 2.2: Integrated luminosity delivered to (green), and recorded by ATLAS (yellow) during stable beams and for 7 TeV center-of-mass energy in 2010 (a) and 2011 (b) as a function of time.

In 2010, the maximum instantaneous luminosity was 2×10^{32} cm^{-2}s^{-1} and the bunch crossing interval was not smaller than 150 ns. Therefore, the out-of-time pile-up effect was small. The data taking conditions in 2011 were more challenging than those of 2010, since the protons bunch crossing interval was reduced to 50 ns and the instantaneous luminosity reached values of 3×10^{33} cm^{-2}s^{-1}. The average number of pp collisions per bunch crossing was between 3 and 8 until summer 2011, with a global average for this period of $<\mu>\approx 6$. In the second semester of 2011, $<\mu>$ increased to reach values between 5 and 17, with a global average of $<\mu>\approx 12$. Figure 2.3 shows a Z boson candidate event decaying into two muons with 20 reconstructed vertices, which was recorded on September 14th, 2011.

2.2 The ATLAS detector

ATLAS (A Toroidal LHC ApparatuS) is one of the experiments located in the LHC ring. It was built as a general purpose detector to observe/measure the particles generated in the LHC pp collisions at unprecedented energies and luminosity. The detector was installed in the underground cavern between 2003 and 2008 after many years of research, preparation and construction. It has a length of 44 m and a height of 25 m. Its layout is shown in Figure 2.4. ATLAS has a wide physics program aiming basically at: Standard Model precision measurements, searches for the Higgs boson and for physics beyond the Standard Model. As a collaboration, ATLAS consists of more than 2900 physicists and engineers from 172 different institutions around the world.

Like many collider experiments, ATLAS is built of several sub-detectors, configured in concentric layers around the interaction point, each one serving a different purpose. From the interaction point outwards, the first subdetector is a tracking one: the *inner detector* which is used for tracking of charged particles. The subsequent subdetector is the *calorimeter system*, which measures the energy of particles by total absorption. It is divided into an electromagnetic (in charge of measuring the energy and position of photons and electrons with high precision) and a hadronic component (where hadrons and other particles with enough energy to not be fully absorbed in the electromagnetic calorimeters deposit their remaining energy). Of all known interacting particles, only muons and neutrinos are penetrating enough to reach beyond the calorimeters. Neutrinos interact only weakly and escape the detector without leaving traces. A

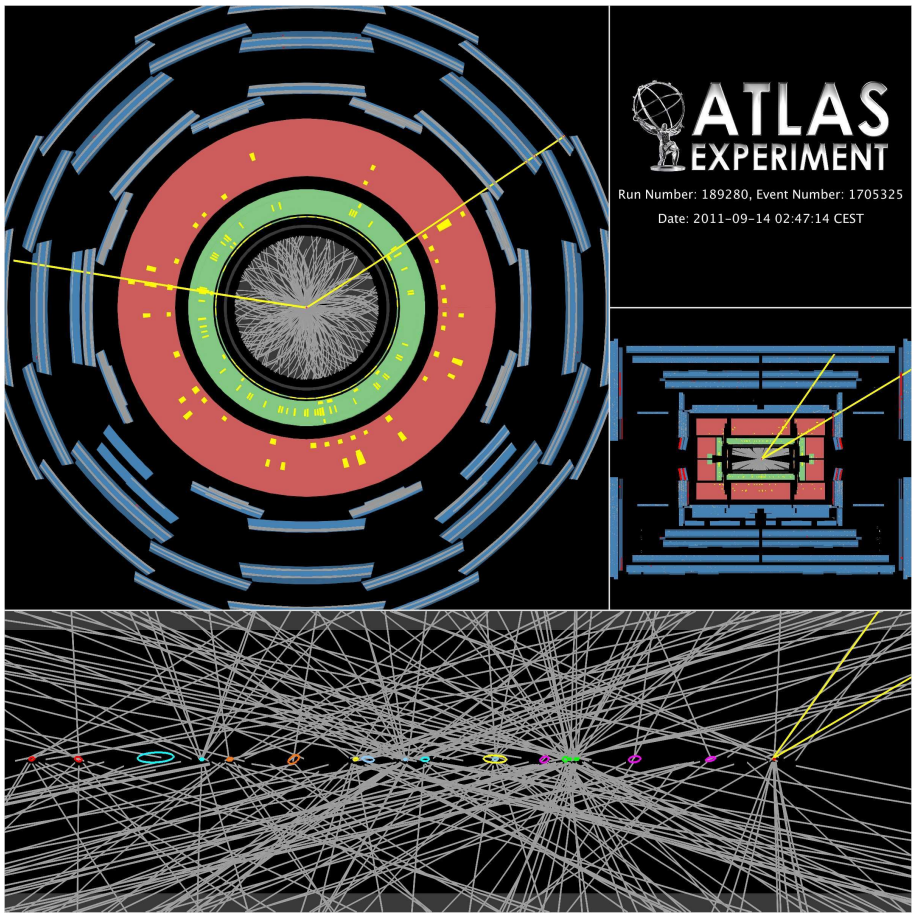


Figure 2.3: Z boson candidate event with 20 reconstructed vertices, with the Z boson decaying into two muons. A track p_{T} threshold of 0.4 GeV was used. All tracks are required to have at least 2 pixel and 7 SCT hits. The reconstructed vertex error elipses are shown scaled up by a factor of 20 so that they are visible.

tracking sub-detector is placed outside of the calorimeters to measure the muon's momentum: the *muon spectrometer*. In this way the detection of all the particles is achieved (except for neutrinos), see Figure 2.5. Both tracking sub-detectors (the inner detector and the muon spectrometer) operate in a magnetic field, provided by a solenoidal and a toroidal magnet system respectively. Additional detectors exist to monitor the luminosity and the beam position.

The ATLAS coordinate system and the sub-detectors most relevant for the analysis presented in this thesis are described below. The *trigger system*, which helps selecting interesting events to be stored and analyzed is described afterwards. Finally, the *Monte Carlo simulation* of the ATLAS detector is briefly discussed, as it is an important tool to understand how the detector operates and has a central role in the analyses presented in later chapters.

ATLAS coordinate system

The ATLAS coordinate system is a cartesian right-handed coordinate system, with the nominal collision point at the origin. The $x-$axis points to the center of the LHC ring, the $z-$axis follows the beam direction and the $y-$axis points upwards. The azimuthal angle ϕ is the angle with the positive $x-$axis in the $x-y$ plane. The polar angle θ is defined as the angle with the positive $z-$axis. The distance in the $x-y$ plane is defined as $r=\sqrt{x^2+y^2}$. An important and

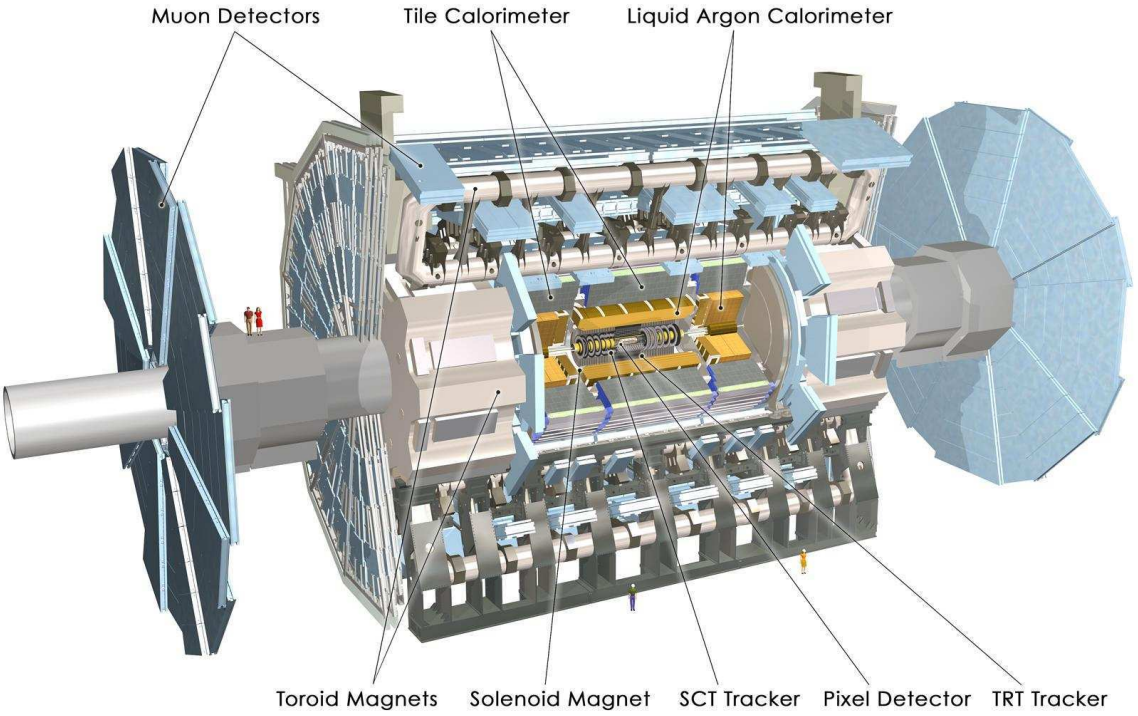


Figure 2.4: A detailed computer-generated image of the ATLAS detector and its subsystems [52]

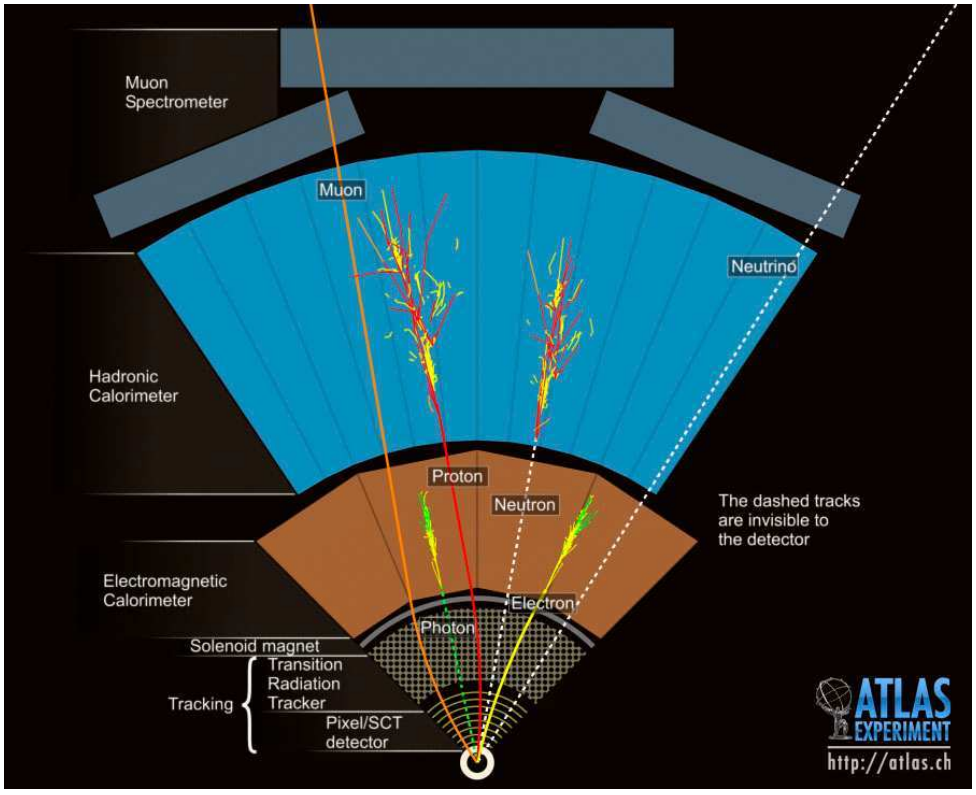


Figure 2.5: Illustration of particle detection in the subsystems of the ATLAS detector.

very used variable is the pseudorapidity which is defined as $\eta = -\ln[\tan(\theta/2)]$. The rapidity is defined as $y = 0.5 \times \ln[(E + p_z)/(E - p_z)]$, where E denotes the energy and p_z is the component of the momentum along the $z-$axis. For massless objects, the rapidity and pseudorapidity are equal. The variables ϕ and η define the pseudorapidity-azimuthal angle space ($\eta - \phi$ space) and parametrize the different directions in which particles are emitted/detected with respect to the interaction point. In this space the distance is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The pseudorapidity also helps in the description of the general features of the detector. In the following, the central detector region (generally up to $|\eta| < 1.6$) is referred to as *barrel*, the more forward region (up to $|\eta| < 2.5$ unless otherwise noted) as *endcap* and the forwardmost pseudorapidities are simply called *forward* region.

In a collision the overall boost along the $z-$axis is not known since the partons that give rise to a given process carry an unknown fraction of the proton momentum. This is why mostly boost-invariant quantities are used in most analyses. Differences in pseudorapidity and rapidity are invariant under boosts along the z axis, as well as the transverse momentum (p_T, projection of the particle momentum on the $x - y$ plane).

2.2.1 The inner detector

The Inner Detector (ID) is the ATLAS subdetector closest to the beam axis, where the density of particles is the largest. High granularity and good radiation tolerance are required. Its purposes are to reconstruct the trajectories of charged particles traversing it, to identify vertices and to measure the momenta of charged particles. Since the ID is embedded in a 2 Tesla magnetic field, the reconstruction of the trajectory allows momentum determination of charged particles. A more detailed description of the ATLAS magnet can be found in Section 2.2.4.

With a total length of 7024 mm, an inner radius of 45 mm and a outer radius of 1150 mm, the ID is as close as 10 mm to the beam pipe. It consists of three tracking devices: the *pixel detector*, the *Semi-Conductor Tracker* (SCT) and the *Transition Radiation Tracker* (TRT). Their configuration is illustrated in Figure 2.6. Each of the three sub-detectors is divided into a barrel region, where the detector modules are laid out in cylindrical layers, and an end-cap region, where disks are used to increase the detector coverage in η (up to ± 2.5) without a large increase in detector surface.

2.2.1.1 The pixel detector

The pixel detector is the innermost part of the inner detector, consisting of three concentric layers around the beam axis and three disks perpendicular to the beam axis on each side of the interaction point. It provides an η coverage up to 2.5 and a complete ϕ coverage. It has a resolution of 10 μm in the $r - \phi$ direction and 115 μm in the z direction. The detector consists of 1744 modules (external dimensions 19×63 mm^2). Each module is built of a 250 μm thick layer of silicon connected to a read-out electronic. The size of a pixel is typically 50 $\mu m \times 400$ μm. The total number of pixels per module is thus 47232. The read-out is performed using 46080 different channels. Electron-hole pairs are created in the silicon when charged particles pass through (in a number which is proportional to the particle's energy loss) and an externally applied electric field can be used to collect that charge and generate a signal out of the doped silicon used. The layout of the pixel detector primarily allows determining the particle's position and identifying the primary and secondary vertices. Secondarily, the energy deposited in combination with a measurement of the particle trajectory can be interpreted in order to identify the particle that traversed the detector [58].

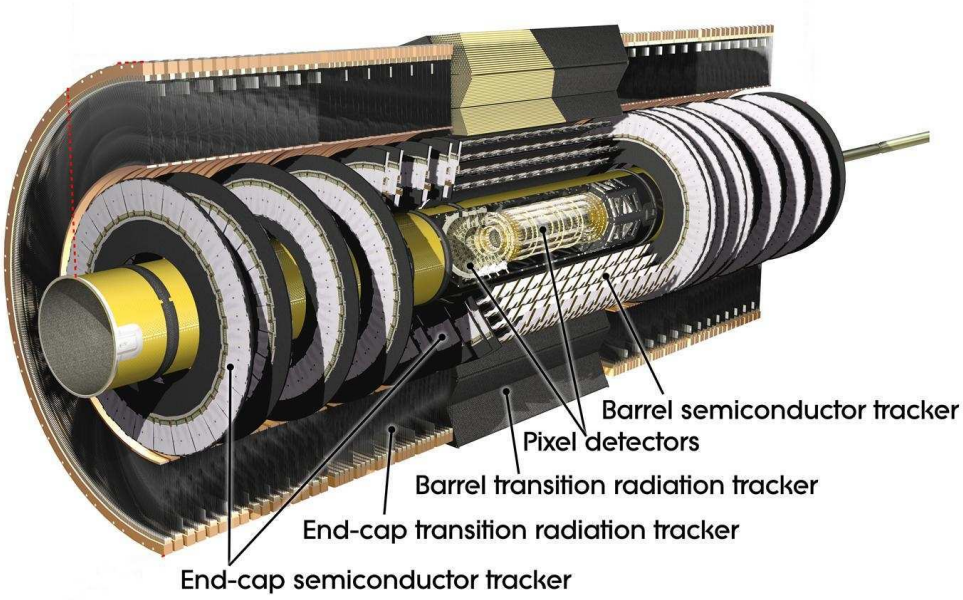


Figure 2.6: Computer-generated image of the ATLAS inner detector system [52]

2.2.1.2 The semi-conductor tracker

The middle component of the inner detector is the Semi-Conductor Tracker (SCT), which is composed of four coaxial cylindrical layers in the barrel part around the beam axis and nine endcap disks along the beam line on each side. It provides an η coverage up to 2.5 and a complete ϕ coverage. It has a resolution of 17 μm in the $r-\phi$ direction and 580 μm in the z direction. The detector consists of 15912 silicon strip sensors mounted on 4088 modules of rectangular shape in the barrel or trapezoidal shape in the end-cap. Its detection principle is similar to that of the pixel detector, although the lower particle density allows using silicon strips rather than small rectangular pixels. The strips are configured in two layers under a small angle with respect to each other, such that a position measurement along the strip length can be obtained from hits in overlapping strips.

2.2.1.3 The transition radiation tracker

The outer component of the inner detector is the Transition Radiation Tracker (TRT). It uses 4 mm diameter gaseous straw tubes that are interleaved with transition radiation material. The barrel contains 73 such layers (the straws are parallel to the beam axis and are 144 cm long), while 160 layers of 37 cm long straws are used to build each of the two end-cap modules on the plane transverse to the beam direction. It has an η coverage up to 2.0 and only provides $r-\phi$ information. It has a resolution in the $r-\phi$ direction of 130 μm per straw. Particles passing through the tubes ionize the gas inside the tube, and the charged atoms and electrons are pulled apart by the electric field existing between the walls of the tube and the thin wire going through the center of the tube. The detected charge is used to build a signal that allows determining if there was a hit, and how far from the center of the straw it occurred (since the drift time is

converted into the distance of the track to the wire).

The TRT also helps in the discrimination between electrons and pions. As was mentioned, the straws are interleaved with foils and fibers of a material with a different refraction index than that of vacuum. When any particle goes through these fibers, it radiates energy in the form of low-energy (∼keV) photons (transition radiation). The amount of energy radiated is proportional to $\gamma = E/m$, where E is the energy of the particle and m is its mass. For a given energy, electrons thus radiate about 250 times more energy than pions. The gas mixture inside the straw tubes contains xenon to increase the transition radiation emitted and produce stronger signals. This signal will have a higher amplitude when an electron passes through than when a pion does it.

Summarizing, a particle originating from the interaction point, given that it satisfies $|\eta| < 2.0$ and $p_{\mathrm{T}} > 5.0$ GeV, typically gives rise to three pixel hits, four SCT hits and around 30 TRT hits, allowing a precise reconstruction of the particle trajectory. Eventhough the semiconductor trackers provide three-dimensional space points with high precision, it is the high number of TRT hits over the larger part of the track length which contributes the most to the momentum measurement. In addition, electron identification is provided by the TRT. The semiconductor trackers also allow impact parameter measurements and vertexing for heavy-flavour and τ-lepton tagging. The secondary vertex measurement performance is enhanced by the innermost layer of pixels.

2.2.2 The calorimeter system

After the inner detector we find the calorimeter system, which is situated outside the solenoidal magnet that surrounds the inner detector. It extends from approximately 1.4 m to 4.2 m from the interaction point in the transverse plane, with a complete ϕ coverage and an η coverage up to 4.9. Its length is 12.20 m. Its purpose is to measure the energy of all interacting particles (except muons, which are highly penetrating and are not absorbed in the calorimeters). Neutrinos interact with very low probability and remain undetected. The transverse component of the undetected energy, known as the *missing transverse energy*, can nevertheless be estimated by means of the expected energy balance in the transverse plane. The performance of the calorimeters has a direct influence on this quantity, which will be discussed in more detail in Chapter 4. The configuration of the calorimeters is depicted in Figure 2.7.

The ATLAS calorimeter system uses sampling technology, i.e. layers of passive, dense material with high stopping power alternated with layers of active material. Particle cascades (showers) are produced through interactions of the incident particle in both the absorber and the active material. Particles deposit their energy in both the absorber and the active material, eventhough only the energy deposited in the active material is measured. The incident particle interacts with the calorimeter material through several processes [59]:

- Photons interact with matter basically by three processes: the photoelectric effect, Compton scattering, and pair production. All of these contribute at varying levels over different energy regimes. For high energy photons, pair production is the dominant process, while for very low energy photons, the photoelectric effect is the most probable interaction.

- Positrons and electrons can ionize atoms or molecules, emit Cerenkov radiation or produce bremsstrahlung radiation. Except for special Cerenkov detectors, the ionization and bremsstrahlung processes are the most important. Ionization is more important in lower energy regimes. While for energies greater than about 100 MeV, bremsstrahlung is the dominant process for electrons and positrons to lose energy.

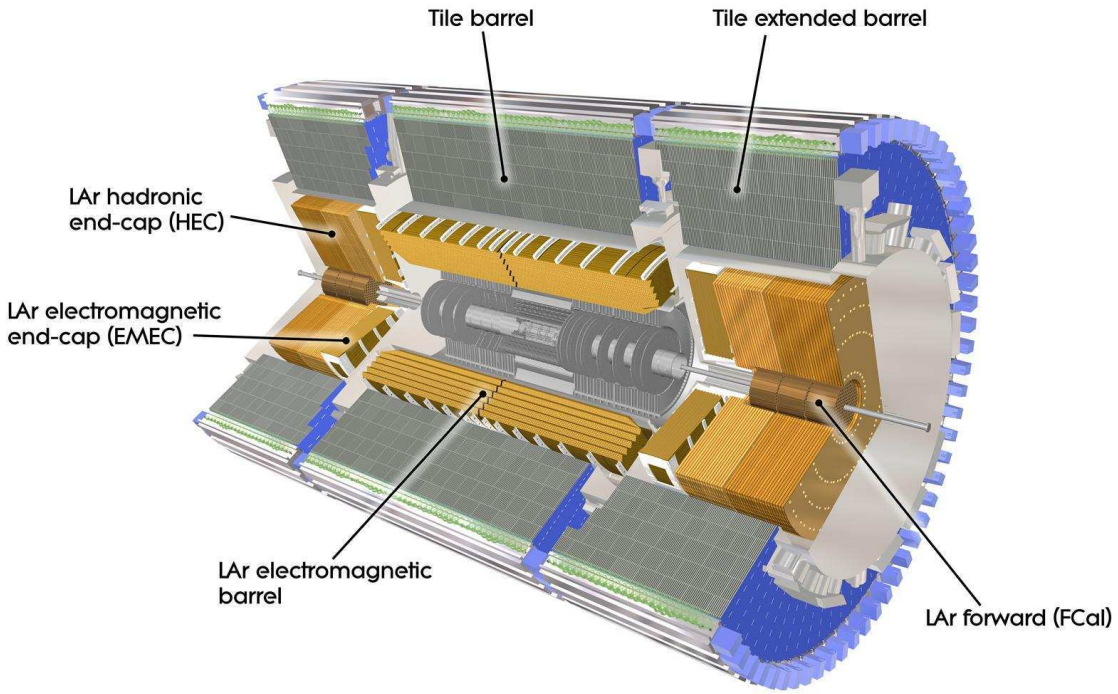


Figure 2.7: Computer-generated image of the ATLAS calorimeter system [52]

Due to the interactions of electrons, positrons, and photons described above, a rapid proliferation of secondary particles occurs when any of these particles travels through matter. The particle cascade produced is known as *electromagnetic shower*.

- Hadrons can lose their energy in many different ways, for example via ionization (if the hadron is charged) or through nuclear interactions. The result is a proliferation of secondary mesons, nucleons, etc., which will further interact in the material, resulting in a *hadronic shower*.

  Hadronic showers are longer and less denser than electromagnetic ones. In any hadronic shower there are significant amounts of energy deposited electromagnetically by hadrons such as π^0s and ηs through decays to $\gamma\gamma$ (*EM energy*). In addition, there is a component of energy in a hadronic shower that is absorbed in nuclear breakups and excitations which is fundamentally undetectable in the calorimeter (*invisible energy*). There may also be some particles, mostly neutrinos, which will leave the detector without being detected, and the energy loss that they represent is called *escaped energy*. The average ratio between signals from the electromagnetic and hadronic component of the shower depend on the choice of passive and active materials and their relative thickness. In the case of the ATLAS calorimeter, the electromagnetic component of the shower is detected more efficiently than the hadronic component. This is known as non compensation effect in calorimetry and is directly related with impossibility to account for the invisible energy in hadronic showers. There are also compensating calorimeters, with a response for electromagnetic shower component similar to the one for the hadronic one.

If the number of successive layers is large, the primary particle will lose almost all of its initial energy. But this energy is not completely measured by the calorimeter system even with an

optimal design, due mainly to energy that remains in the absorber, in non-instrumented regions of the detector and to the non compensation of the calorimeter. For this reason, calibrations need to be applied to properly measure the energy of the incident particle (see Section 4.7.2).

One important characteristic of calorimeters is their resolution. Since calorimetry is based on statistical processes, the measurement accuracy improves with increasing energy. The accuracy, called calorimeter resolution, improve as $\sigma_E/E = a/\sqrt{E}$. This expression is often expanded when discussing calorimeter performance on analysis objects to account for noise in the electronics, energy that might be lost in non-instrumented areas of the detector and various other calibration effects like mis-calibration, mis-alignment, non-compensation, etc.:

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C, \tag{2.3}$$

where $\oplus$ is used to indicate addition in quadrature and S, N and C are the stochastic, noise and constant terms, respectively. The second term accounts for instrumental effects and its relative contribution decreases with E. This component may limit the low-energy performance of calorimeters. The third component is due to calibration errors, non-uniformities and non-linearities in the detector. It sets the limit for the performance at very high energies. Due to the production of neutral pions in the hadronic showers, the visible energy fraction fluctuates heavily from event to event, which results in a worse energy resolution for hadronic showers compared to electromagnetic ones.

ATLAS includes an electromagnetic and a hadronic calorimeter, with different characteristics in order to account for the different properties of electromagnetic and hadronic showers.

2.2.2.1 The electromagnetic calorimeters

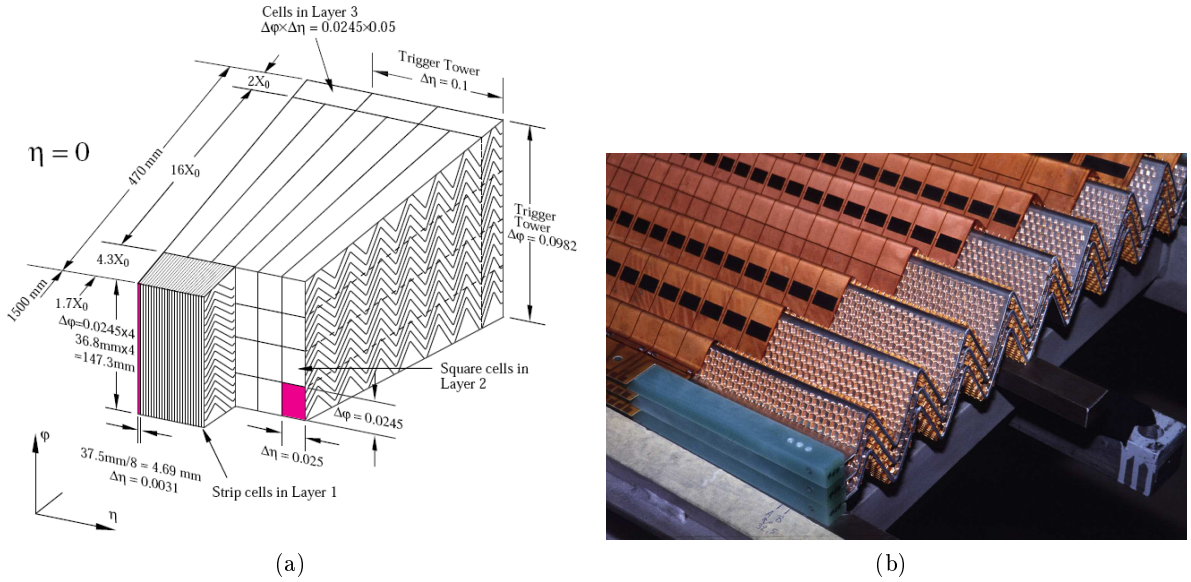
The electromagnetic (EM) calorimeters are designed for the identification and measurement of photon and electron energy and position. They are placed in front of the hadronic calorimeters, in part because particles that interact only electromagnetically require less material to be absorbed. It consists of a barrel that covers $|\eta|$ <1.475 and two coaxial endcap wheels at 1.375< $|\eta|$ <3.2. It uses liquid argon (LAr) as its active material and lead as its absorber arranged into an accordion shape (allowing a finely segmented read-out and providing naturally full ϕ coverage without cracks). The liquid argon is located in the gaps between the 1024 accordion absorbers. Liquid argon was chosen as the active material due to its resistance to radiation and its uniformity, which translates into spatial uniformity in the energy measurement. Charged particles traversing the calorimeter ionize the LAr, and the resulting electrons drift towards the copper electrodes in the read out cells thanks to the presence of an electric field.

The EM calorimeters are located in cryostats, since liquid argon needs to be kept at a temperature of about 88 K. The inner radius of the cryostat is 1385 mm, and its outer radius is 2132 mm. The barrel component shares its cryostat vessel with the solenoid magnet (see Section 2.2.4) in order to minimize the amount of inactive material. Between the barrel and each endcap, around $|\eta| = 1.4$, some space (known as crack region $1.375 < |\eta| < 1.52$) is available for cables and services for the inner detector. Below $|\eta| = 1.8$ the EM calorimeter is complemented by a presampler, inside the cryostat, that consists of a thin (11/5 mm in the barrel/end-cap) layer of instrumented LAr to provide additional measurement information. Conceptually, the presampler acts as the sampling layer for the material before the calorimeter, which would be the absorber in this case. Scintillator slabs are also placed in the gap between barrel and endcap ($1.0 < |\eta| < 1.6$).

The characteristic interaction distance of an electromagnetic shower is the radiation length X_0 of the material, defined as the mean distance over which an electron loses all but $1/e$ of its

energy. The total number of radiation lengths traversed by a particle in the electromagnetic calorimeter ranges from $22X_0$ to $33X_0$.

The modules in the LAr calorimeter barrel are divided into three layers (see Figure 2.8). The first layer, finely segmented in strips of $\Delta\eta \times \Delta\phi$ =0.0031×0.098, is used to reconstruct the η position of electromagnetic showers and provides information on particle identification. The first layer is approximately $4X_0$ thick only, and thus only a small fraction of the energy of a particle is deposited there. The second layer is the one that collects the largest fraction of the energy deposited by the shower. Its length is of about $17X_0$. The third layer (about $1X_0$-$10X_0$ thick) only collects the shower tail. The electrodes in the second and third layer are grouped in towers of $\Delta\eta \times \Delta\phi$ =0.0245×0.025 and 0.0245×0.05 respectively. The first wheel of the LAr endcaps is segmented in three layers with a granularity that varies as a function of the pseudorapidity and of the layer, while the second wheel (for larger pseudorapidities) is segmented in two sections in depth and has a coarser granularity (0.1×0.1) in both layers (see Ref. [52]).



(a) (b)

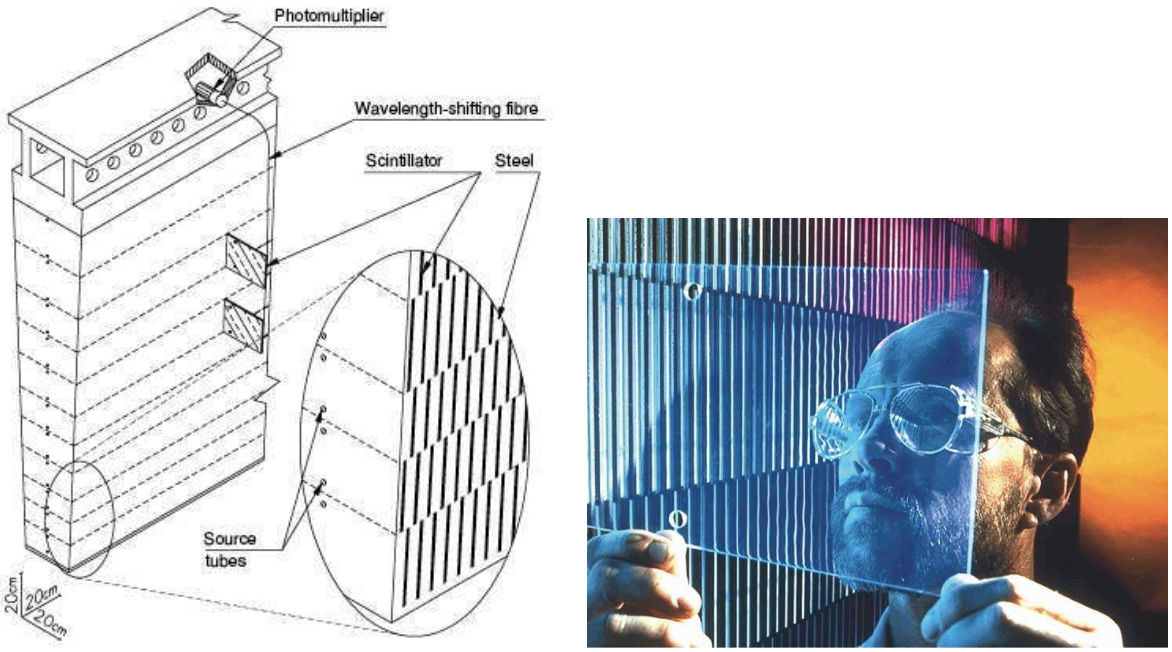
Figure 2.8: View of a module in the electromagnetic calorimeter, showing the typical accordion shape and the granularity of the different layers [52].

The drift time in the LAr under the influence of the 2 kV electric field is approximately 450 ns (which is much longer than the nominal LHC bunch spacing, and impacts the energy reconstruction in the calorimeters during operation in nominal conditions). Since the drift velocity enters in the energy measurement, the detector conditions that influence it, such as high voltage and liquid argon temperature and density, need to be continuously kept under control to reduce the fluctuations in the energy measurement.

2.2.2.2 The hadronic calorimeters

Hadronic showers are the result of nuclear interactions and develop over distances larger that electromagnetic ones. The depth of material required to contain them is expressed in terms of the nuclear interaction length λ of the passive material. Hadrons deposit some fraction of their energy in the EM calorimeters, but they are not fully absorbed. Then, the role of the hadronic calorimeters is to contain the showers of high-energy hadrons and it is achieved through two different calorimeters: the Hadronic Barrel and the Hadronic End-Cap.

The hadronic barrel calorimeter The hadronic barrel calorimeter, also known as the Tile calorimeter, uses scintillating plastic tiles as its active material and steel as its absorber. It consists of one barrel module (covering $0 < |\eta| < 0.8$) and two extended barrel modules (covering $0.8 < |\eta| < 1.7$). Its inner and outer radii are 2280 and 4230 mm, respectively. The gap between them is filled with cables, services and power supplies for the Inner Detector and the EM Calorimeter. For this reason, scintillating tiles have been placed there to act as the active material for the non-instrumented services. The barrel module surrounds the EM barrel calorimeter, while the extended barrel modules surround the end-cap cryostats housing the EM end-cap, hadronic end-cap and forward calorimeters. The barrel and extended barrels are divided azimuthally into 64 modules. The modules are installed radially around the beam axis to get almost full coverage in ϕ. The 3 mm tiles are placed perpendicular to the colliding beams, as shown in Figure 2.9. Particles interact with the active material producing scintillation light proportional to the deposited energy, which is collected by wavelength shifting fibers. Readout cells are built by grouping fibers together in projective towers in η, and the scintillation light is collected by photomultiplier tubes at each end of the tiles, where the signal is amplified and detected.



(a) (b)

Figure 2.9: Structure of a module of the Tile calorimeter, showing the placement of the tiles and the readout components [52].

As for the EM calorimeter the read-out is segmented longitudinally. In this case in three layers in both barrel and extended barrel. The cell granularity is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for the first two layers and $\Delta\eta \times \Delta\phi = 0.1 \times 0.2$ for the third one. The thickness of the hadronic calorimeter is approximately 10λ.

The hadronic end-cap calorimeter The Hadronic End-Cap (HEC) calorimeter uses liquid argon as its active material and copper as its absorber. It is located directly behind the end-cap electromagnetic calorimeter and shares the same LAr cryostats. With an η coverage of $1.5 < |\eta| < 3.2$ it is placed at 2.03 m from the interaction point. Each hadronic

end-cap consists of two wheels where 8.5 mm active gaps are sandwiched between copper plates of different widths (25 and 50 mm for inner and outer wheel respectively). The gaps are split into four regions of 1.8 mm each by three electrodes. The middle electrode serves as the read-out electrode and defines the $\eta \times \phi$ segmentation (of 0.1×0.1 for $|\eta| < 2.5$ and 0.2×0.2 elsewhere), while the other two have a voltage of 1800 V applied to them. Each wheel is built from 32 identical wedge-shaped modules, assembled with fixtures at the periphery and at the central bore. The HEC read-out is segmented along the direction of the shower in four read-out layers. The HEC calorimeter shares the read-out service routing and infrastructure with the EM calorimeter.

2.2.2.3 The forward calorimeter

A large pseudorapidity coverage is necessary because it allows an improvement in the estimation of p_{T} of neutrinos or other particles that could escape the detector, through the improvement in the measure of forward jets. The coverage in the very forward region, $3.1 < |\eta| < 4.9$, is provided by the Forward Calorimeter (FCAL), constitute by three wheels on either side: one electromagnetic component and two hadronic components. At high rapidity, the proximity to the beam and the limited amount of space determine the design of the forward calorimeter. To allow for shower containment, the absorber must be dense (the electromagnetic component uses copper, while the two hadronic components employ tungsten). The gaps must be narrow to avoid ion build-up and have a fast readout time. Each wheel consists of a metal matrix, with regularly spaced longitudinal channels filled with the electrode structure consisting of concentric rods and tubes parallel to the beam axis. The LAr in the gap between the rod and the tube is the sensitive medium. Like the HEC, it shares many of its read-out and environmental features with the EM calorimeter. The FCAL is approximately 10 interaction lengths deep. The distribution of material upstream of the different calorimeters layers in terms of interaction lengths is shown in Figure 2.10.

2.2.3 The muon spectrometer

The Muon Spectrometer (MS) is the outermost ATLAS subdetector. Its purpose is to trigger muons as well as to reconstruct their trajectory and calculate their momentum. The first is achieved using coarse chambers: the Resistive Plate Chambers (RPC) and the Thin Gap Chambers (TGC), while chambers with higher precision are employed for precise muon measurements: the Monitored Drift Tube (MDT) chambers and the Cathode Strip Chambers (CSC). The chambers provide coverage up to $|\eta| = 2.7$ (except at $|\eta| \approx 0$, where all the services from the other detectors are routed out of ATLAS). It extends from the end of the calorimeters, about 5 meters away from the beam, to about 10 meters radially. Along the direction of the beam, they extend from around 7 m to over 20 m in each direction and occupies a volume of around 16000 m^3. The arrangement is such that a particle originating from the interaction point will traverse three layers of muon stations as it is bended by a toroidal magnetic field (see Section 2.2.4). A computer-generated image of the muon spectrometer is shown in Figure 2.11, indicating the four different types of chambers and the magnet toroids.

Since this magnetic field bends particles only on a plane of constant azimuth, ϕ, that passes through the beam axis, no precise ϕ information is needed to reconstruct the muon momentum. Then, the precision muon chambers are built to measure the coordinate of interest, η. The precision chambers have a long charge collection time, which results in a long read-out latency. This latency does not guarantee that the signals can be used for triggering purposes. For this reason, the precision chambers are complemented by trigger chambers.

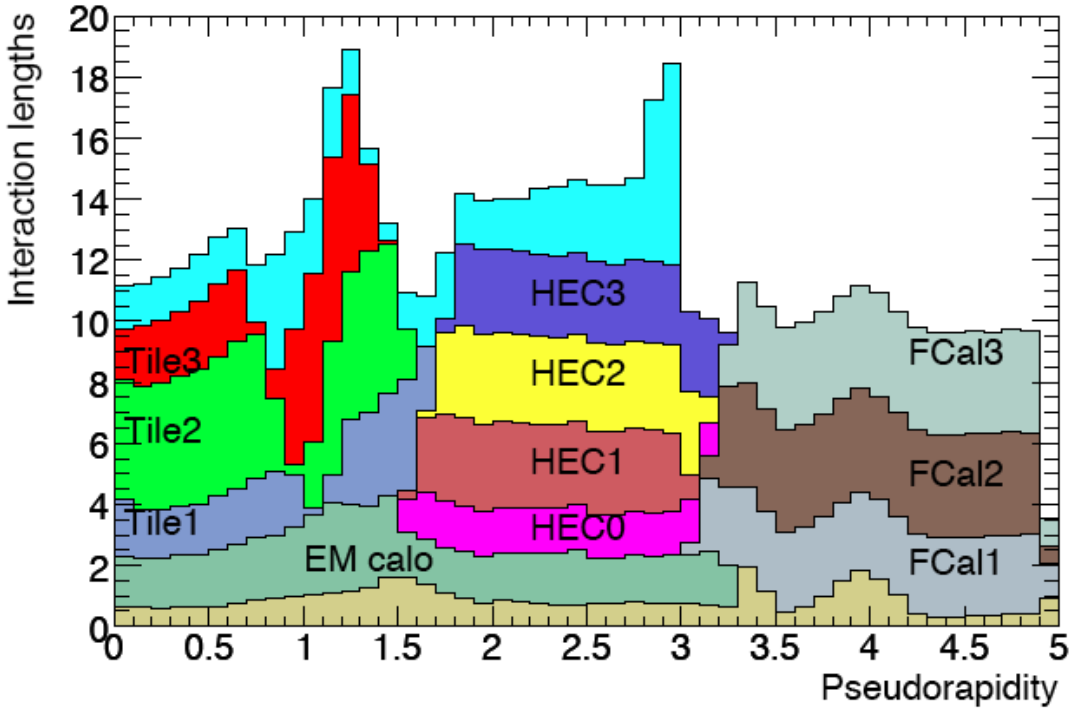


Figure 2.10: Cumulative amount of material in units of interaction length in each layer of the ATLAS calorimeters. The total amount of material in front of the first active layer of the muon spectrometer (up to $|\eta| < 3.0$) is also shown in cyan [52].

The resolution on the transverse momentum p_{T} achieved by the muon spectrometer is $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 10\%$ at $p_{\mathrm{T}} = 1$ TeV. Their momentum reconstruction resolution is best at around 100 GeV (3%) and increases to about 4% at low momenta due to fluctuations in the energy lost by muons in the calorimeters. This resolution can be improved to about 2% if tracks measurements from the muon spectrometer are combined with those from the inner detector. The actual reconstruction of muon tracks is further described in Chapter 4.

The Monitored Drift Tube (MDT) Chambers consist of two multilayers of aluminium pressurized drift tubes of about 3 cm in diameter, using an argon/CO_2 mixture as the drift gas. A tungsten-rhenium anode wire in the tube collects the charge produced through the ionization caused by the muons traversing the tube. The measured drift time is used to determine the coordinate of the muon with the required precision. The tubes are placed into multilayers in order to improve the tracking performance. MDTs do not operate properly in very high particle density environments. For this reason, they are not placed close to the beam.

The Cathode Strip Chambers (CSC) are multiwire proportional chambers. The wires are at a voltage of 1900 V. They lie sandwiched between strip cathodes in a gas mixture (argon/CO_2), creating an electric field that can be used to measure charge created by an ionizing particle. The coordinates of a traversing charged particle are obtained from the relative measurement of induced charge on adjacent cathode strips. The strips on each of the two cathode planes are positioned orthogonally, thus allowing for the determination of two coordinates: in the bending direction and in ϕ.

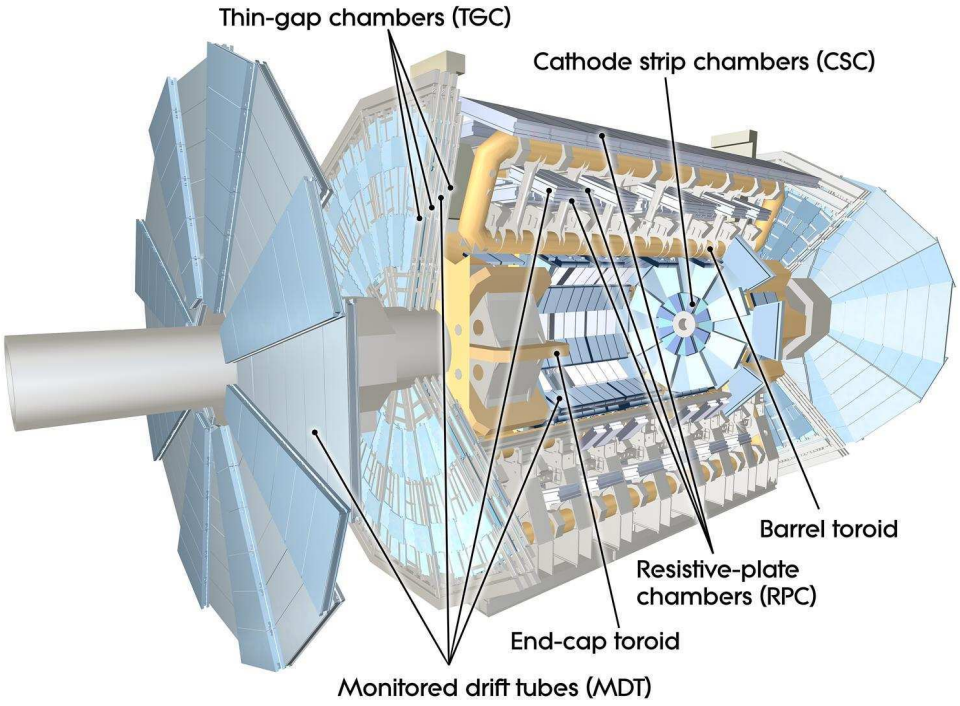


Figure 2.11: Computer-generated image of the muon spectrometer. The different types of chambers are labeled. The magnet toroids are colored in orange [52].

The Resistive Plate Chambers (RPC) consist of two rectangular detectors. Each one being a gaseous detector where the gas mixture (primarily $C_2H_2F_4$ is used) is enclosed by two resistive rectangular plates separated by 2 mm. A voltage of 9.8 kV is applied between the rectangular plates (leading to a drift time of about 5 ns), such that the charge created by the ionizing particle are leaded towards the anode plane, where the signal is read out. The RPCs are built up of two detectors to provide redundancy that reduces the noise and other backgrounds, such as photons and low-energy neutrons in the cavern. RPCs are placed on the same support structure than some of the barrel MDTs.

The Thin Gap Chambers (TGC): the technology used by the TGCs is similar to the one used by the RPCs. The gas is a mixture of CO_2 and n-C_5H_{12}. The distance between the wires is 1.8 mm, while the distance from the wire to the cathode is 1.4 mm. The drift time combined with the signal propagation time in the electrodes guarantee that the signals arrive to the read-out system within 25 ns (the LHC bunch separation). The TGCs are built of two or three gaseous detector planes. These trigger chambers measure both coordinates of the track, one in the bending (η) plane and one in the non-bending (ϕ) plane. TGCs, unlike the RPCs, have their own support structure.

2.2.3.1 The muon spectrometer barrel

The muon spectrometer barrel consists of three concentric cylindrical layers of muon stations with a pseudorapidity coverage up to $|\eta| = 1$ and almost complete in ϕ. The stations are organized in sixteen sectors, thus following the structure of the eight barrel toroid magnet coils (see Section 2.2.4). Long and short chambers are used in each sector with an up and down radial displacement that provides chamber overlaps for full ϕ coverage. Muon stations in the innermost

layer are single MDT chambers located just outside the hadronic calorimeter and named Barrel Inner (BI) chambers. Stations in the middle layer consist of a Barrel Middle (BM) MDT chamber with a RPC on either side and are situated inside the barrel toroid magnet. The outer layer consists of stations that each comprise a Barrel Outer (BO) MDT chamber and a RPC positioned just outside the barrel toroid magnet.

2.2.3.2 The muon spectrometer end-caps

The two endcaps consist of four disks each covering a pseudorapidity range of $1.0 < |\eta| < 2.7$. The three main wheels are placed 7 m, 13 m and 21 m away from the interaction point along the beam direction. A smaller wheel is placed on top of the end-cap toroid at a distance of 11 m from the interaction point. This guarantee that three tracking points are available in the full η coverage without having to increase the radial size of the last wheel significantly. Most part of the disks consist of trapezoidally shaped MDT chambers, except the first end-cap wheel, in the region $2.0 < |\eta| < 2.7$, where particle density is higher, so CSCs are used instead to get a better spacial resolution. The trigger information for high η regions ($1.05 < |\eta| < 2.4$) is provided by four TGC layers. The first one placed on the inner side of the first end-cap wheel, while the other three are placed in the proximity of the third wheel, behind the end-cap toroid magnet cryostat. One of those three is placed in front of the MDTs of the third wheel, while the other two are behind the MDTs, a few centimeters apart.

2.2.4 The magnet system

The ATLAS magnet system generates a magnetic field configuration that bends the particle's trajectories when traversing the inner detector and the muon spectrometer. It is 22 m long in diameter and 26 m long in length, with a stored energy of 1.6 GJ and provides the magnetic field over a volume of approximately 12000 m^3. It consists of two superconducting magnet systems: a toroidal one which provides a magnetic field inside the muon spectrometer and a central solenoid which provides it inside the inner detector. The curvature of the charged particle's trajectory when passing through the magnetic field can be used to determine its momentum. Figure 2.12 shows the ATLAS Magnet System layout.

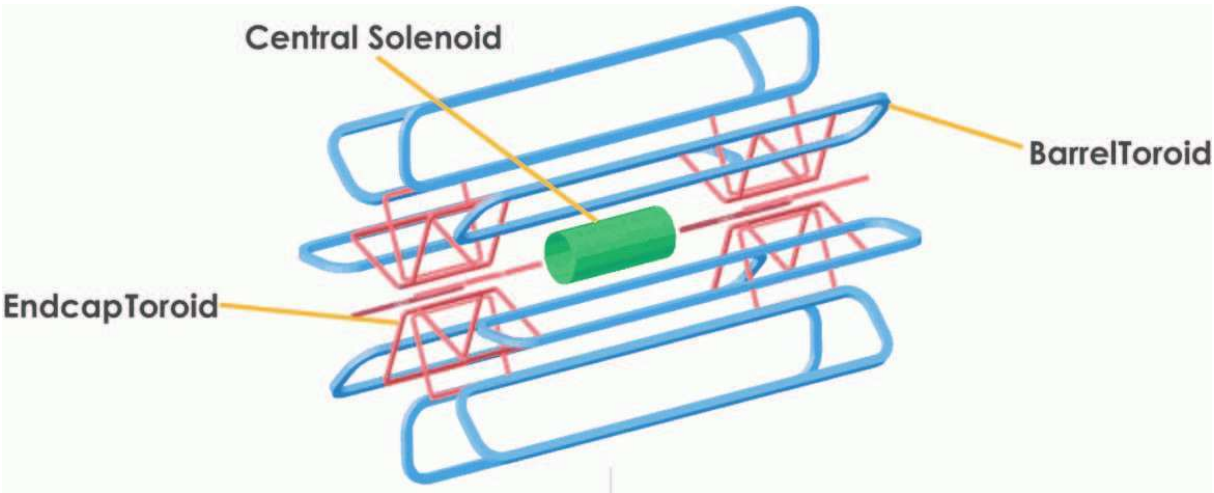


Figure 2.12: Computer-generated image of the ATLAS Magnet System. The eight barrel toroid coils, with the end-cap coils interleaved are visible. The solenoid winding lies inside the calorimeter volume.

2.2.4.1 The central solenoid

It is aligned on the beam axis and provides a 2 T axial magnetic field for the inner detector. It lies inside the cryostat used for the barrel EM calorimeter and in front of the calorimeters, to reduce the amount of material in front of them. The design had to cope with stringent constraints on the amount of material, to have a small impact on the energy measurement in the calorimeters. It has an axial length of 5.8 m and a diameter of about 2.5 m. It is cooled using liquid helium down to a temperature of 4.5 K. Its nominal operating current is 7730 A. This provides the 2-Tesla magnetic field that embeds the ID. This field is returned through the steel support structure of the hadronic barrel calorimeter, beyond the calorimeter volume. In order to monitor the magnetic field inside the inner detector, four Nuclear Magnetic Resonance (NMR) probes are glued to the cryostat wall, equally spaced in ϕ, measuring the magnitude of the magnetic field with an accuracy of 10 μT [60].

2.2.4.2 The toroidal magnet system

The toroidal magnet system provides bending power for the muon spectrometer. It is built up of a barrel toroid and two endcap toroids. The barrel toroid consists of eight superconducting rectangular coils, each encased in a cryostat. The total assembly weighs 830 tons. It has an inner and outer diameters of 9.4 m and 20.1 m respectively, and an axial length of 23.5 m. Cooling down to the nominal operational temperature of 4.6 K takes 5 weeks. The field strength provided by the barrel toroid at the nominal operational current of 20.5 kA varies from 0.15 T to 2.5 T.

The endcap toroid systems consist of eight coils each, which are located interleaved with the barrel toroid coils on either side, thus generating a magnetic field in the endcap regions of the muon spectrometer. With an inner and outer diameter of 1.65 m and 10.7 m and an axial length of 5.0 m, each endcap toroid weighs 239 tons. Powered in series with the barrel toroid, the endcap toroids generate a field strength that varies from 0.2 T to 0.35 T at nominal operational current.

2.2.5 The forward detectors

The ATLAS detector is complemented by other detectors, allowing additional measurements for physics or monitoring purposes. In the forward region five extra detectors are located outside the ATLAS detector volume. Their location along the LHC beam axis are shown in Figure 2.13. The main purpose of the first two is to determine the luminosity delivered to ATLAS: LUCID (LUminosity measurement using Cerenkov Integrating Detector) and ALFA (Absolute Luminosity For ATLAS) [61].

LUCID consists of 32 aluminum tubes located around the beam at a distance of 17 m from the interaction point on each side of the detector. The tubes are 1.5 m long and 15 mm in diameter, and point towards the interaction point to minimize the signals created by particles coming from other interactions in the cavern. The tubes are filled with C_4F_{10} gas where Cherenkov light is created and collected/amplified by photomultipliers. The signal from the photomultipliers is amplified and a discriminator is used to determine if a tube was hit. Each hit is associated to a given bunch crossing and this information is sent to the ATLAS trigger and the data acquisition systems. LUCID has been designed to work from luminosities of 10^{27} cm^{-2}s^{-1} up to 4×10^{33} cm^{-2}s^{-1}. The luminosity measurement is done bunch by bunch.

ALFA consists of 8 scintillating fiber detectors located in Romans Pots on each side of ATLAS detector at 240 m distance from the interaction point, above and below the LHC beam axis. ALFA is designed to approach as close as 1 mm to the beam. The ALFA detector was installed about the middle of 2011. Its goal is to measure in dedicated runs of low luminosity (from 10^{27}

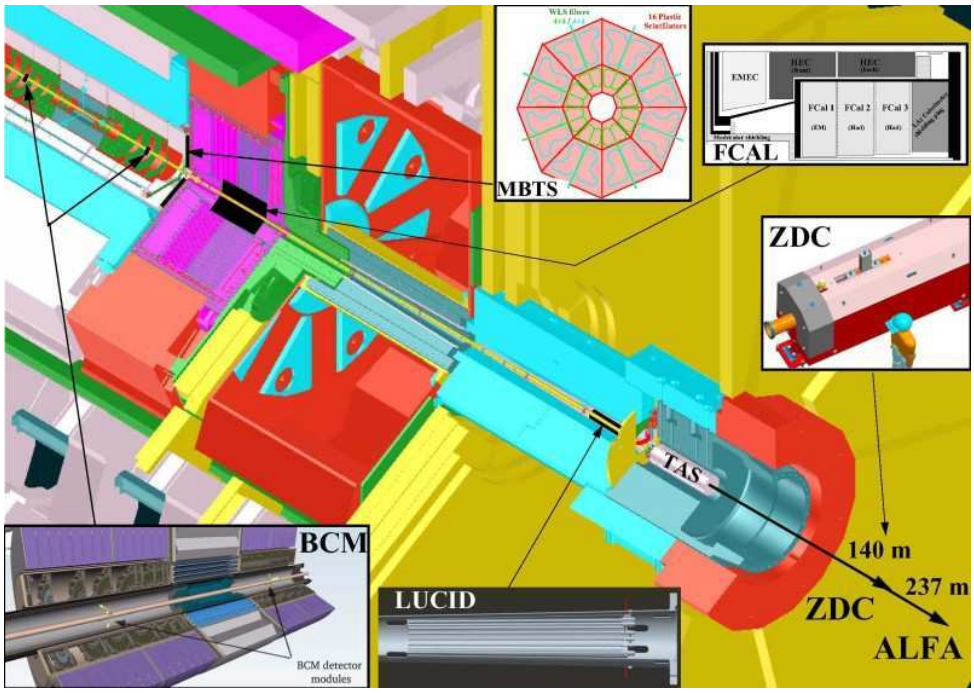


Figure 2.13: Location of the different ATLAS monitoring detectors along the LHC beam axis.

to 10^{28} cm^{-2}s^{-1}) the total pp cross section and absolute luminosity thus providing a calibration point for LUCID. It is expected that ALFA will measure the absolute luminosity for ATLAS with an accuracy of about 3% in 2012.

The third system is the Minimum Bias Trigger Scintillators (MBTS). It has been designed to detect and trigger minimum bias[2] activity during a bunch crossing in the early operation of ATLAS, but it is also used for relative luminosity measurement. The MBTS consists of 32 2 cm thick scintillator paddles assembled on two disks. The disks are located in front of each end-cap calorimeter, placed at 3.5 m away from each side of the interaction point. They cover a pseudorapidity range of $2.09 < |\eta| < 3.84$. Wavelength-shifting fibers collect the light produced in each scintillator and transport it to the photomultiplier tubes for signal amplification. The signals are then sent to the central trigger processor.

The fourth system is the Beam Condition Monitor (BCM). It has the main function of monitoring the beam against beam losses that could cause detector damages, but it is also used as luminosity monitor.The BCM consists of two sets of diamond sensors located 184 cm away from the interaction point in the direction of the beam and 5.5 cm away in the radial distance. Each side has four modules with two sensors each. The sensors are 1 cm×1 cm in size and 500 μm thick, and operate at a voltage of 1000 V. If large beam losses are detected, BCM sends a signal to the LHC, which causes an abort and a controlled emptying of the accelerator ring. Its signal is also sent to the ATLAS detector in order to take the necessary actions to minimize the damage.

The fifth detector is the Zero-Degree Calorimeter (ZDC). It plays a key role in determining the centrality of heavy-ion collisions and is used as luminosity monitor in pp collisions. It is located at ±140 m from the interaction point, just beyond the point where the common straight-section vacuum-pipe divides back into two independent beam-pipes. The ZDC modules consist of layers of alternating quartz rods and tungsten plates which will measure neutral particles at

[2]i.e. partonic interactions with transverse momenta too small for perturbation theory to be valid

Subdetector	Required resolution	η coverage
Inner Detector	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	± 2.5
Electromagnetic Calorimeter	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2
Hadronic Calorimeter		
Barrel and Endcap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2
Forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$
Muon Spectrometer	$\sigma_{p_T}/p_T = 10\%$ at p_T =1 TeV	± 2.7

Figure 2.14: Performance requirements for the subdetectors of the ATLAS detector [52]

pseudorapidities $|\eta| > 8.2$. Its estimated lifetime at 10^{34} cm^{-2}s^{-1} is of a few months. The measurements from these luminosity monitors are complemented by measurements in the ID to provide the final luminosity estimate. The final normalization of the luminosity is based on Van der Meer scans, which determine the size of the colliding beams [62]. The luminosity measurement used in this thesis was performed using the LUCID detector, while the other detectors were used to establish the long-term stability of the LUCID measurement.

ATLAS main performance goals in terms of resolution as well as the acceptance of each subdetector are summarized in Table 2.14.

2.2.6 The trigger system

The LHC interaction rate is getting higher and higher. In nominal conditions the bunch crossing frequency will be 40 MHz, with about 23 interactions occurring per bunch crossing, leading to approximately one billion events per second. Not all these events are interesting as a big part of them correspond to minimum bias events. Moreover, the current capabilities for recording events offline (recording space and storage space) are not enough. Therefore, the ATLAS trigger system has been developed and implemented in order to select interesting collision events within a minimal time. The trigger system uses simple criteria to rapidly decide which events in a particle detector to keep when only a small fraction of the total can be recorded. The event rejection process needs to reject the background without biasing the selection of the physics signals. This is done in three subsequent levels: the *level-1 trigger*, the *level-2 trigger* and the *event filter*. The idea is that each level selects the data that becomes an input for the following, which has more time available and more information to take a better decision.

2.2.6.1 The level-1 trigger

The level-1 trigger is designed to accept as much as 75000 events per second (i.e resulting in a reduction of the event rate from 40 MHz to 75 kHz in nominal conditions). The selection is made using information of reduced granularity from the calorimeters and the trigger muon chambers. This allows for selections based on the presence of muons, electrons, photons, taus, jets, event transverse energy and missing transverse energy. The MBTS is also used in the level-1 trigger to select events where a collision actually happened. The level-1 trigger defines Regions of Interest (RoIs) in the (η, ϕ) space where object candidates satisfy a required energy threshold, and based on the required multiplicity of RoIs, the event is passed or not to the level-2 trigger. In order to minimize the propagation time through cables, the dedicated electronics are located as close as possible to the ATLAS detector. The trigger decision time is constrained to 2.5 μs to guarantee that the event is still stored in the front-end buffers and, thus, can be sent further along the read-out chain.

2.2.6.2 The level-2 trigger

The level-2 trigger and the event filter are part of the software trigger subsystem called High Level Trigger (HLT). The level-2 trigger is designed to reduce the event rate to 3.5 kHz and has an average latency of 40 ms. It analyzes the level-1 RoIs with more complex object reconstruction algorithms and using information from all detectors, including the inner detector. If the event is accepted by the level-2 trigger, it is sent to the event filter for further analysis.

2.2.6.3 The event filter

The event filter (EF) is designed to reduce the event rate from 3.5 kHz to the 200-400 Hz, which is the rate at which ATLAS events can be written to disk. It runs a physics reconstruction close to the offline one (looser selection criteria are applied by the event filter with respect to the criteria using offline) in about 4 seconds/event (with an event size of approximately 1.3 Mbyte) and assigns each event into *streams* based on the triggers passed. The event streams are built separating events of interest for different analyses in different datasets. This mean that the same event can appear in different streams depending on the stream definition. They are defined to minimize this overlap. As an example we can find electron, muon, photon, jet, minimum bias, etc. streams.

One important concept in the trigger system is the *prescaling*. The prescale determines how often an event that passed a given trigger is accepted. A prescale of 1 means that all events selected by the trigger are accepted, while a prescale of 1000 means that events passing the trigger will only be accepted one out of a thousand times. Pre-scales are needed to control the rate at which commonly occurring events are recorded. The sequence of algorithms that defines a certain trigger object at each trigger level will be referred to as a *trigger chain*. The final energy threshold and quality requirements are naturally determined by the last trigger level used. The naming convention for a trigger chain is:

[LEVEL][N][TYPE(S)][THRESHOLD][ISOLATION][QUALITY],

where LEVEL refers to the trigger level used, TYPE(S) specifies the object(s) candidate(s) (i.e electron, muon, photon, jet, etc. or a combination of them), N indicates its multiplicity, THRESHOLD is a number corresponding to a transverse momentum (or energy) threshold applied, ISOLATION indicates the object isolation and QUALITY refers to the severity of requirements in the algorithm. For instance, a trigger chain called *EF_e20_medium* triggers electrons at the Event Filter level with a transverse momentum larger than about 20 GeV and which satisfies the loose requirements defined in the reconstruction algorithm used at the event filter level. The individual decisions can also be logically combined to more complex trigger items.

The available trigger chains are defined in terms of a *trigger menu*. A trigger menu consists of a set of trigger chains with their corresponding prescales. It is chosen taking into account the LHC luminosity and the physics program for each data taking period. The events that pass any trigger chain in the given trigger menu are arranged in *luminosity blocks* (LB). A LB is a time unit that corresponds typically to a couple of minutes of data taking. The beam condition, detector performance and trigger configuration are stored for each LB and can be later used for analysis tasks needing time-dependent information (for example the luminosity calculation or the exclusion of luminosity blocks with problems in the detector). A *run* is the period between the start and the end of the data taking and it is divided in LBs. In ideal conditions, a run should start at the beginning of a LHC fill and end when the beam is dumped, but due to problems in the data acquisition system we usually have several runs by LHC fill. Runs are grouped into *periods*.

Runs in the same period share the same general detector conditions, machine configuration and trigger menu. Periods are usually denoted by letters (i.e. period A, period B...).

To summarize, the trigger pipeline is the following: when particles interact with the detector components, signals are generated in the detector front-end electronics, these signals are buffered in the front-end electronics and transmitted to the off-detector electronics (outside the experimental hall or "cavern" where ATLAS lies) if an accept signal is received from the level-1 trigger. Then the data are sent to the read-out drivers and after packaging and processing to the read-out servers, where they stay until the decision from the level-2 trigger has been taken. If the event passes the level-2 trigger menu the requested information is sent to the event builders that pack the data in one data structure. This data structure is then moved to the EF. If the event passes the EF menu it is written to disks where it can be accessed for analysis. This data is structured in runs, which at the same time are divided into LBs. The trigger chain used in the different analyses presented in this thesis are discussed at the beginning of each analysis chapter.

Chapter 3

ATLAS Event Simulation

Contents

3.1 Overview

In real life, the LHC produces events that are stored by the data acquisition system of ATLAS (see Chapter 2). In simulated reality, Monte Carlo event generators play the role of the LHC machine and detector simulation programs play the role of the ATLAS detector. During the preparation phase of an experiment, simulation provides the environment to develop and understand the detector, to develop analysis strategies, to estimate the sensitivity to different physics processes, to develop and validate object reconstruction algorithms (see Chapter 4), to optimize the trigger menus, and so on. During the running phase of an experiment, simulation is used to compare predictions of theoretical models against the real data.

In ATLAS, the event simulation and reconstruction is performed in the Athena framework [63]. This software is in charge of the following processes:

Event generation: corresponds to the phase of proton-proton (pp) collision events generation. It takes care of the production and decay of particles in a given process. Several event generators are available. Further description is provided in Section 3.2.

Detector simulation: is the simulation of interactions between the generated particles and the detector.

Digitization: corresponds to the simulation of the detector readout, i.e. the conversion of energy deposited in the detector to times, currents and voltages for readout electronics. The output format of the simulation is identical to the real detector output format. The simulation of the ATLAS detector response is briefly discussed in Section 3.3.

Reconstruction: in this step a set of object reconstruction algorithms are applied. These algorithms are applied to both simulation and real data in exactly the same way. This phase is fully described in Chapter 4.

3.2 Phenomenology and simulation of pp collisions

Protons are baryons composed of two up quarks and a down quark (known as the valence quarks) held together by the strong force mediated by gluons, as well as additional virtual quark-antiquark pairs (which constitute the sea) that flit in and out of existence as the time passes. For the generation of processes that occur in pp collisions a basic ingredient is the Quantum Chromodynamics (QCD), the theory of strong interactions. QCD was introduced in Chapter 1 Section 1.2.2.2. QCD has been verified in a wealth of fixed-target and collider experiments since 1960s (most relevant results are collected in Ref. [64]). Despite its success as theoretical framework for strong interactions, QCD is a theory where calculations are difficult. Many approximations exist in order to increase QCD's predictive power [65]. A brief overview of the main QCD concepts is presented in the next paragraphs.

3.2.1 Quantum Chromodynamics (QCD): key concepts

3.2.1.1 Asymptotic freedom and confinement

When calculating physical observables from the SM lagrangian, all Feynman diagrams involved in the process under study are summed over. Often this involves divergent contributions from loop diagrams. A renormalization procedure is necessary to treat infinities arising in calculated quantities, to be able to compare the results with the experimental measurements. The price paid in this procedure is the introduction of an arbitrary scale, called renormalization scale μ_0, in order to keep consistent dimensions (units) for all quantities. Imposing the independence of the physical observable from μ_0 reveals that the strong coupling constant, α_S, depends on the energy scale at which one observes the coupling. The strong coupling constant can be written at leading order as:

$$\alpha_S(\mu^2) = \frac{\alpha_S(\mu_0^2)}{1 + (11 - \frac{2}{3}n_f)\frac{\alpha_S(\mu_0^2)}{2\pi} ln\frac{\mu^2}{\mu_0^2}}, \tag{3.1}$$

where n_f is the number of quark flavors (up, down, strange, charm, bottom, and top) in the theory and μ is the variable that express the energy dependence. Two important properties of QCD are direct consequences of the behavior of α_S:

- **Asymptotic freedom:** the value 11 in the denominator of Eq. 3.1 arises from the self-interaction of the gluons and determines that α_S decreases as μ increases. This is known as asymptotic freedom. It means that at very high energies and short distances quarks and gluons interact very weakly, i.e. they behave as free particles. The low value of α_S makes possible to use perturbation theory at high energies. At high energies, the *parton model* [66] can be used to describe hadrons with a good precision, since it treats the partons as free and non interacting.

- **Confinement:** the size of α_S increases at low energies and large distances, which means that the force between partons does not diminish as they are separated. Therefore, partons are tightly bound together in color neutral combinations (hadrons) and rapidly recombine into such combination when forced apart in high energy collisions. It explains why partons can not be observed directly in experiments. The increase of α_S makes perturbative

calculations less and less reliable. Approximations working for specific phase space regions need to be developed. For soft interactions at low energy, the lattice QCD approach can be used, but the numerical calculations are highly CPU consuming. Therefore, models based on experimental measurements are mainly used (see Section 3.2.2.5).

3.2.1.2 Factorization

Factorization [67] refers to the set of theorems that establish that the short distance component of the scattering process described by perturbative QCD can be separated from the non-perturbative long distance component, i.e allows to separate parts that describe the final-state hadron formation from the perturbative hard interaction part among the partonic constituents. Factorization is a byproduct of a procedure that absorbs singularities into physical quantities (in a similar way as renormalization does). For this reason, a new scale called the factorization scale μ_F^2, is introduced.

3.2.2 Monte Carlo event simulation chain

The Monte Carlo event simulation enables the description of high energy collisions from its initial states to the final ones. A large number of events, consisting with a list of final state particles and expected momentum probability densities, are generated. Pseudorandom numbers are used to simulate event-to-event fluctuations intrinsic to quantum processes. MC generators make use of the factorization principle (see Section 3.2.1.2), so the different phases of the proton-proton collision are considered independently. This makes possible to simulate complex final states with hundreds of particles with a reasonable accuracy. These phases are dominated by different dynamics, and the most appropriate techniques can be applied to describe each of them separately.

A simulated event is built up in several steps, as illustrated in Figure 3.1. Not all these steps are relevant in all processes. However, all of them are discussed here to give an idea of the complexity of the event generation process. Initially two protons collide and undergo a deep inelastic interaction, with a large momentum transfer. The two protons can be seen as a bag of partons (quarks and gluons). The collision between two partons from the protons, one from each side, is known as the *hard interaction*. The outcome of this hard interaction is the scattering at large angle of the two partons or their annihilation into a massive particle or resonance. The decays of short-lived resonances produced in pp collisions, e.g. top quarks or Z' bosons are regarded as part of the hard interaction. The hard interaction can be calculated perturbatively.

Since the partons involved in the hard interaction are color (and often electromagnetically) charged, they can radiate gluons. Emission associated with the two incoming colliding partons are called *initial state radiation* (ISR), while emission that can be associated with the partons created by the hard interaction are called *final state radiation* (FSR). The gluons emitted can themselves emit further gluons or produce quark/antiquark pairs leading to the formation of *parton showers*. There are two kind of parton showers:

- **Initial state shower** is one that develops from an incoming parton of the hard interaction.
- **Final state shower** is one that develops from an outgoing parton of the hard interaction.

The radiation process is governed by perturbative QCD. As the event is evolved downwards in momentum scales it ultimately reaches the region, at scales of order 1 GeV, in which the strong coupling constant value increase and perturbation theory breaks down. Therefore at this scale the perturbative evolution must be terminated and replaced by a non-perturbative *hadronisation*

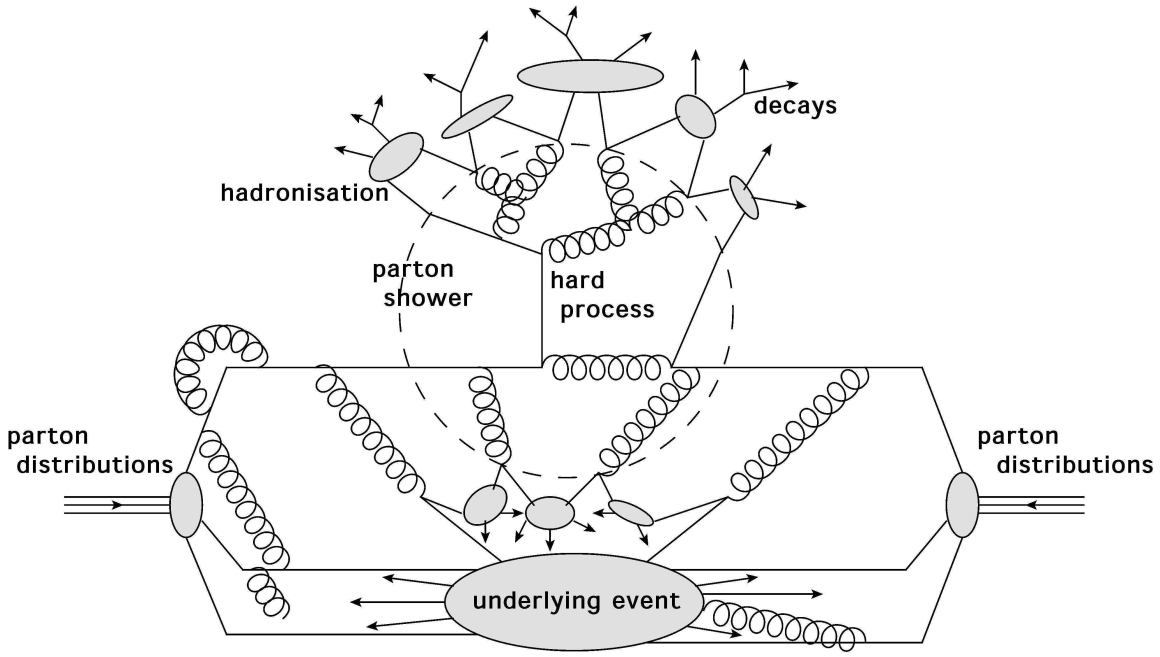


Figure 3.1: General structure of a pp collision [68].

process, in which the partons are bound into colorless hadrons. Many of the hadrons that are produced during hadronization are unstable. Phenomenological models are used to simulate their decay to lighter hadrons that are long-lived enough to be considered stable on the time-scales of particle physics detectors. Then, what one detects instead of quarks and gluons are collimated shower of particles, known as *jets*. Jets are discussed in more details in Chapter 4.

Finally, to complete the picture, we need to understand the evolution of the fragments of the initial interacting hadrons. This evolution cannot be entirely independent of what happens in the hard event, because at least color quantum numbers must be exchanged to guarantee the overall neutrality and conservation of baryon number. The remaining partons of the initial hadrons can also interact with each other. These multiple interactions go on to produce additional partons throughout the event, which may contribute to any observable, in addition to those from the hard process and associated parton showers that we are primarily interested in. This part of the event structure is known as the *underlying event*.

3.2.2.1 Hard interaction: Matrix Element (ME)

The hard interaction involves large momentum transfers. At this scale, the partons behave as asymptotically free, making possible to describe this process using perturbation theory. Let's consider the simple process of pp scattering (as in the LHC case): $pp \to X$ (see Figure 3.2). According to the parton model, the respective cross section can be expressed as [69]:

$$\begin{aligned} \sigma_{pp\to X} &= \sum_{a,b} \int dx_1 dx_2 \int f_a(x_1,\mu_\mathrm{F}) f_b(x_2,\mu_\mathrm{F}) d\hat{\sigma}_{q_a q_b \to X}(x_1,x_2,\mu_0,\mu_F) \\ &= \sum_{a,b} \int dx_1 dx_2 \int d\Phi_X f_a(x_1,\mu_\mathrm{F}) f_b(x_2,\mu_\mathrm{F}) \times \frac{1}{2x_1x_2s} |\mathcal{M}_{q_a q_b \to X}|^2(\Phi_X,\mu_\mathrm{F},\mu_0), \end{aligned} \tag{3.2}$$

where:

- $x_{1,2}$ are the fractions of proton momentum carried by the two incoming partons.
- μ_F and μ_0 are usually set to the characteristic scale of the process Q^2. For example, $Q^2 = M^2$ for the production of an $s-$channel resonance of mass M, while $Q^2 = p_{\mathrm{T}}^2$ for the production of a pair of massless particles with transverse momentum p_{T}.
- $f_{a(b)}(x, \mu_F)$ are the Parton Distribution Functions (PDFs). They describe the probability for a parton with flavor $a(b)$ to carry a fraction x of the total proton momentum, when probed at the μ_F scale. PDFs cannot be calculated, since it depends on the non-perturbative physics of the proton wave function. They are extracted from global fits to data collected at previous collider experiments.
- $\hat{\sigma}_{q_a q_b \to X}$ denotes the parton level cross section for the production of the final state X through the initial partons a and b. It depends on the momentum given by the final state phase space Φ_X, on μ_F and μ_0. The fully parton level cross section is given by the product of the corresponding matrix element square $|\mathcal{M}_{q_a q_b \to X}|^2$ (averaged over initial-state spin and color degrees of freedom) and the parton flux $\frac{1}{2x_1 x_2 s}$, where s is the centre-of-mass energy squared.

Eq. 3.2 implies that perturbation theory can be used to calculate the parton level cross section $\hat{\sigma}_{q_a q_b \to X}$, while the non-perturbative component, PDFs, can be included later to obtain the full theoretical prediction. The matrix element can be written as a sum over Feynman diagrams:

$$\mathcal{M}_{q_a q_b \to X} = \sum_i \mathcal{F}^{(i)}_{q_a q_b \to X}. \tag{3.3}$$

Eq. 3.3 is in principle calculable to all orders. Depending on the level of precision required, the matrix element calculation may be treated at the lowest relevant order of perturbation theory (leading order, LO), or to higher order. At present, event generators do not go beyond next-to-leading order, NLO.

3.2.2.2 Parton shower

The previous section describes the generation of a hard interaction of the type $pp \to X$. But as was mentioned before, the hard interaction involves large momentum transfers and therefore the partons participating in it are violently accelerated. Just as accelerated electric charges emit photons, the accelerated colored partons may emit gluons. These gluons can emit further gluons or/and produce quark-antiquark pairs, leading to parton shower generation.

As a result, a more realistic scenario would be consider a process like $pp \to Xg$ for example. In principle, the parton showers represent higher-order corrections to the hard interaction. However, it is not feasible to calculate these corrections exactly. Instead, an approximation scheme is used, in which the dominant contributions are included at each order. These dominant contributions are associated with collinear parton splitting or soft (low-energy) gluon emission. Most of the available Monte Carlo generators control the coherence of this emission by ordering successive emissions by scales related to their transverse momentum or angle wit respect to the incoming parton direction.

Let's consider the almost collinear splitting of a parton of type i into $j + k$, for example $q \to q + g$. If the $n-$parton differential cross section before splitting is $d\sigma_n$, after splitting it becomes (at the leading order in perturbation theory):

$$d\sigma_{n+1} \approx d\sigma_n \frac{\alpha_{\mathrm{S}}}{2\pi} \frac{d\theta^2}{\theta^2} dz d\phi P_{ji}(z, \phi), \tag{3.4}$$

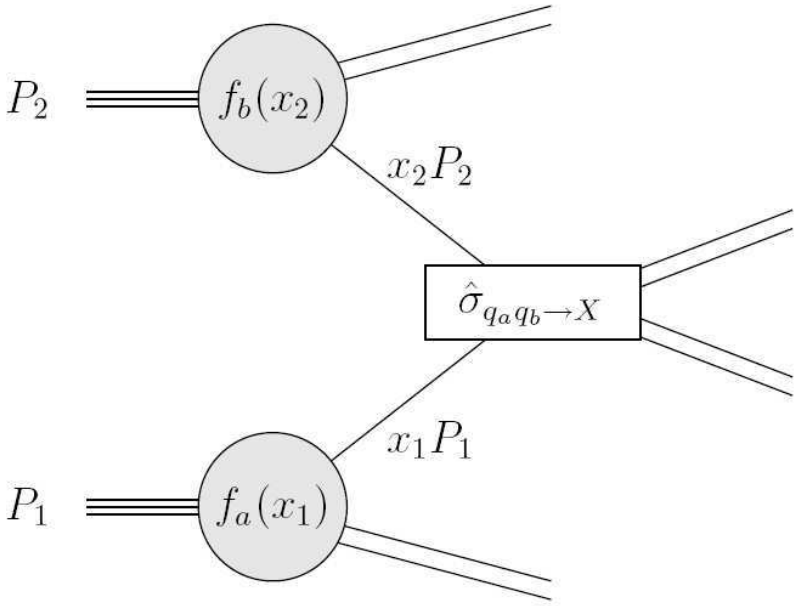


Figure 3.2: Schematic view of the hard interaction in a pp collision: the partons q_a and q_b with momentum fractions x_1 and x_2 of the incoming protons take part in the hard interaction.

where θ and ϕ are the opening angle and azimuthal angle of the splitting and P_{ji} is the $i \to j$ splitting function, which describes the distribution of the fraction z of the energy of i carried by j. A parton shower is generated for each parton of the hard interaction by applying Eq. 3.4 sequentially. A pseudorandom number is used to generate values of z, θ and ϕ for each splitting. An important feature of the showering algorithm, not apparent from Eq. 3.4, is the *evolution variable*. The simplest evolution variable is the virtual mass-squared q^2 (also known as the virtuality) of the partons in the shower, which determines the starting point and the end of a parton shower. The shower starts when $q^2 \approx Q^2$. The shower is terminated when the virtualities have fallen to low values $q^2 = Q_0^2 \approx 1$ GeV^2, where Q_0^2 is called hadronization scale.

The virtual contributions to parton emissions are included in the probability of not splitting during evolution from scale q_1^2 to q_2^2, which is given by the *Sudakov form factor*:

$$\Delta_i(q_1^2, q_2^2) = e^{-\int_{q_2^2}^{q_1^2} \frac{dq^2}{q^2} \frac{\alpha_S}{2\pi} \int_{Q_0^2/q^2}^{1-Q_0^2/q^2} dz \int_0^{2\pi} d\phi P_{ji}(z,\phi)}, \tag{3.5}$$

Eq. 3.5 specifies the range of z in which the splitting is resolvable. An emission that would lie outside this range is too soft or at too small an angle to be detected: it is declared unresolvable and is not included in the shower. At each splitting, the variables z and ϕ are chosen according to the distribution $P_{ji}(z, \phi)$ using the pseudorandom numbers, with z in the resolvable region specified by the limits of integration in the Sudakov form factor.

Final state showers In this case the evolution of the shower proceeds as described above: the primary parton starts at a high energy and a high virtuality scale Q^2 set by the hard interaction, and it looses energy and virtuality until it and all its descendant partons have fallen to a given scale Q_0^2 at which splitting is terminated. At this point the final configuration of parton momentum can be passed to one of the hadronization models described below.

Initial state showers In this case the incoming parton starts at a high energy and low virtuality

and evolves to more virtual masses and lower energies by successive small-angle emissions. The showering of these parton terminates when they collide to initiate the hard interaction, when $q^2 = Q^2$. The partons emitted in the initial state showers may initiate secondary showers that evolve in the same way as final state showers.

The sequence of initial state showering described above is not suitable for Monte Carlo event generation. For example, to generate the following process $q\bar{q} \rightarrow Z^0$, then q and $\bar{q}$ at the end of their initial state showering must have precisely the right 4-momentum to combine to form a system with the Z^0 mass. This will not be the case if their momentum are chosen according to the PDF at the initial low scale. A better procedure, used by most of the event generators, is the *backward evolution*. First the momentum fractions x_1 and x_2 of the incoming partons are chosen using the PDFs at the high hard interaction scale. PDFs have been measured at lower energies and are evolved to higher scales using the QCD evolution equations for parton densities (DGLAP [70, 71, 72]), obtained by averaging Eq. 3.4 over ϕ. The incoming partons are then evolved backwards, gaining energy in each emission. The virtualities and momentum of the incoming partons follow from momentum conservation at the successive splittings in the showers. The only complication is that the no-splitting probability is no longer given by the Sudakov form factor in Eq. 3.5 alone, but rather by that factor modified by a ratio of PDFs at the new, higher, value of x that the parton may evolve back to and its current value:

$$\Delta_i'(q_1^2, q_2^2) = \Delta_i(q_1^2, q_2^2) \frac{f_i(x, q_2^2)}{f_i(x, q_1^2)}. \tag{3.6}$$

According to Eq.3.6, if the parton is in a region in which the PDF decreases rapidly with increasing x, its non-emission probability will be close to one, i.e. its emission probability will be small, and it is more likely that the parton came straight out of the hadron rather than having been produced by evolution of a higher$-x$ parton [65, 68].

3.2.2.3 Combining matrix element and parton showers

Fixed-order matrix elements and parton showers have different advantages and disadvantages. Matrix elements are good to simulate well separated hard partons, but have problems when trying to describe collinear and soft partons and when the final states include a large number of partons. On the other hand, when using parton showers hard emissions are poorly approximated, while soft and collinear parton emissions are well described even for many partons. They should be combined to get the best of both, for an optimal description of multi-parton states. The merging can be done in different ways. One of them consist in supplement the matrix element with Sudakov form factors to give exclusive final states that can be combined with a parton shower. Since those exclusive final states may be already produced by the parton shower, a mechanism to avoid this double counting is needed [15]. There are two main methods used to avoid the double counting: CKKW [73] and MLM matching [74].

The MLM method slices the phase space using geometry and energy conditions, such that soft and collinear emissions are assigned to the parton shower and the rest of the phase space is generated by the matrix element. Then, it applies a jet algorithm to the showered part of the event and identifies all jets with $p_{\mathrm{T}} > Q_{\mathrm{merge}}$, where Q_{merge} is called the merging scale (≈ 15 GeV). If each jet corresponds to one of the parton generated by the matrix element (i.e. is nearby in angle, $\Delta R = 0.7$) and there are no extra jets above scale Q_{merge}, then the event is accepted. Otherwise, it is rejected. Solving the double counting issue ensures that the hard jets always come just from the matrix element. The MLM method also ensures that matched jets

above Q_{merge} have distributions given by the matrix element calculations. The CKKW matching procedure performs a similar slicing of the phase space, but instead of the matching/veto steps of the MLM method, uses an analytical calculation of the Sudakov form factors to veto parton showers.

Combining fixed-order matrix elements with parton showers is a very active research topic, and is important for giving reliable predictions.

Choices for PDFs PDFs play a central role in event generators for the simulation of hard processes and parton showers. The choice of PDF set therefore will influence both cross sections and event shapes. Comparison of experimental data from different processes indicates that PDFs are universal, i.e they do not depend on the physics process. There are several PDFs parametrizations available.The CTEQ[1] and MRST/MSTW[2] collaborations have been especially diligent in regularly presenting updated PDFs. These and others are available in the LHAPDF library[3]. Figure 3.3 shows the behavior of $xf_{a,b}(x, Q^2)$ for each parton flavor obtained by the MSTW group version 2008 NLO 68 [75] as a function of the momentum fraction x for two different Q^2 values.

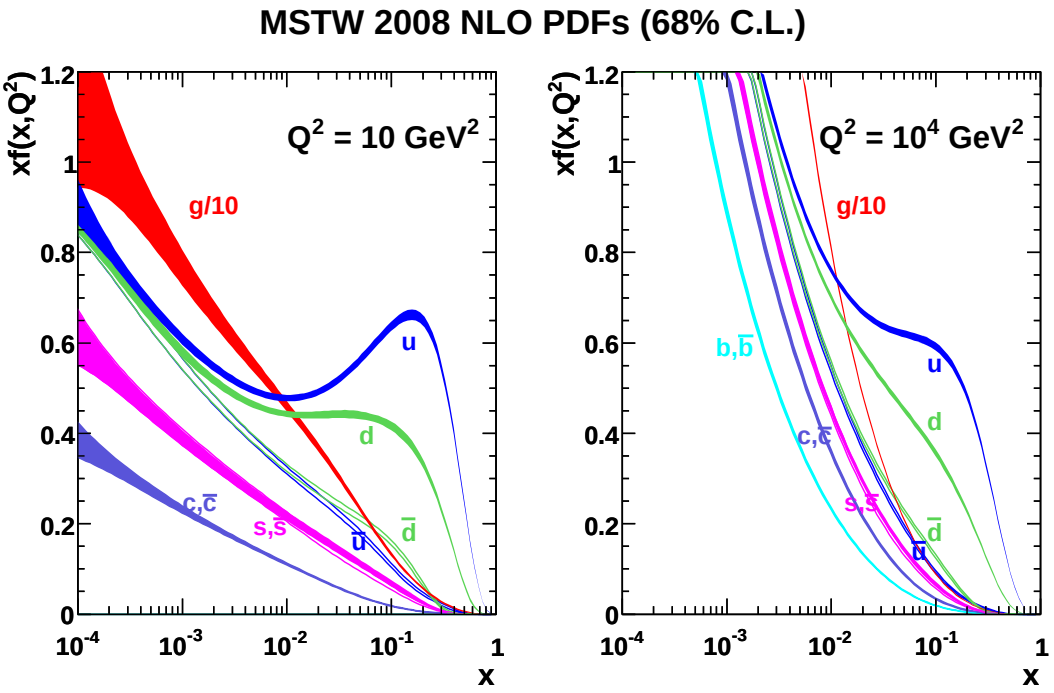


Figure 3.3: Parton distribution functions (PDF) from the MSTW group version 2008 NLO 68 as a function of the momentum fraction x for $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2.

3.2.2.4 Underlying event (UE)

In events containing a hard interaction, the underlying event represents the additional activity which is not directly associated with that main interaction. The UE interactions are soft, so

[1] CTEQ: http://www.phys.psu.edu/ cteq/
[2] MRST/MSTW: http://mstwpdf.hepforge.org/
[3] Les Houches Accord PDFs (LHAPDF): http://lhapdf.hepforge.org/

perturbative approaches don't hold and phenomenological models are used to model them. Parameters in each model are determined using collider data experimental constraints. The UE impacts how the partons from the hard scatter shower and hadronize, and may add additional energy in the direction of the jets from the hard interaction. A good understanding of the UE is needed to be able to measure correctly the properties of the hard interaction. The UE is composed by several phenomena:

Multiple parton interactions (MPI): several parton pairs can undergo into (semi-)hard interactions in an event. The main feature of MPIs is the production of back-to-back jet pairs, with little total p_{T}. For comparison, jets from ISR tend to be aligned with the direction of their parent parton. MPIs are color and kinematically-connected with the rest of the event.

Beam remnants: each incoming particle may leave behind a beam remnant (part of the hadron that does not take active part in the ISR, hard scatter nor MPIs). The beam remnant is modeled using phenomenological models. The basic principles are to keep the color-connection between the beam remnant and the rest of the event and the momentum conservation in the event.

Pile-up: at high luminosity, more than one proton pair may interact per bunch crossing, this is known as pile-up. In order to simulate the pile-up, n semi-hard interactions (Poisson distributed around $< n >$) are generated, with a separated collision vertex. Then, they are added to the hard scatter event.

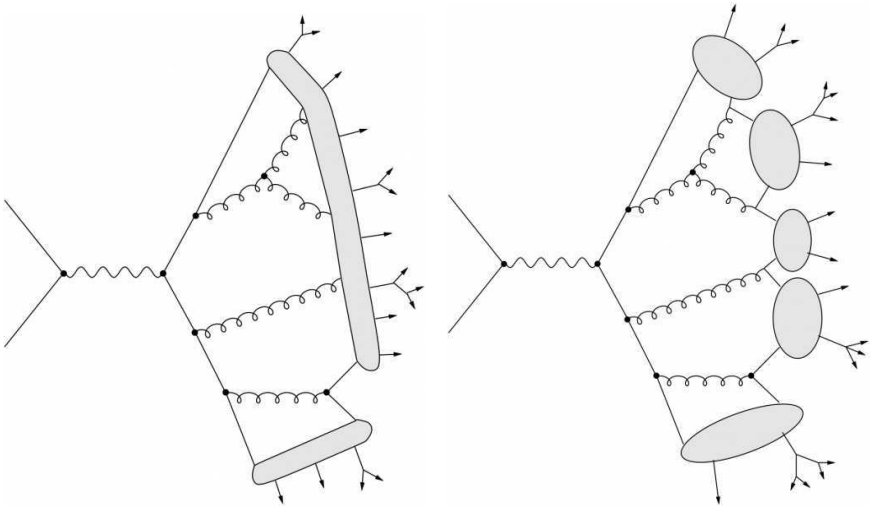
3.2.2.5 Hadronization and decays

Once the low virtuality scale is reached, the phase of hadronization starts. The hadron formation happens at energy scales at which α_{S} is large and perturbation theory can not be used. Several models to simulate the hadronization phase exist. To a good approximation, they are universal (i.e., the hadronization of a given colored system is independent of how that system was produced, so that once tuned on one data set the models are predictive for new collision types or energies). They map effectively partons to hadrons. Some of these hadrons are unstable and decay to lighter and longer-lived particles, which are then used as input to the detector simulation.

The two leading approaches used to model the hadronization process are:

Lund string model [76]: uses as starting point the assumption of linear confinement between partons. This model represents the color force between partons as a string. The two extremities of the string are the opposite color charges. If the string is extended too much, it will break and create two additional color charges at the endpoints: production of a new quark/antiquark pair. The quark and antiquark from two adjacent broken strings can combine to form a hadron (see Figure 3.4).

Cluster model [77]: is based on the color pre-confinement property of the angular-ordered parton shower. After the parton showering, all outgoing gluons are split into light quark-antiquark pairs. Clusters are then formed from color-connected pairs. The high mass clusters split into low mass clusters using string-like mechanism. Cluster splitting continues until all hadrons are build (see Figure 3.4). When combined with angular-ordered parton showers, the cluster model gives a fairly good overall description of high-energy collider data, usually slightly less good than the string model but with fewer parameters.



(a) (b)

Figure 3.4: Schematic view of the results obtained after applying the (a) string model and (b) the cluster model in the hadronization phase of the Monte Carlo generation.

3.2.3 Monte Carlo generators

There are two kinds of Monte Carlo generators:

- **Multi-purpose Monte Carlo generators** which handle all the generation steps described in previous section.
- **Specialized Monte Carlo generators** which handle only individual steps.

The kinematic distributions for a given process may differ between different Monte Carlo generators. Depending on the problem under study, one generator may be more suitable than others. The comparison between different generators is always encouraged. The Monte Carlo generators used in this thesis are listed below, together with their main characteristics.

PYTHIA **[78]** is a multi-purpose Monte Carlo generator for event simulation in pp, e^+e^- and ep colliders. PYTHIA simulates non-diffractive proton-proton collisions using a $2 \to n$ ($n \leq 3$) matrix element at LO to model the hard subprocess, and uses $p_{\mathrm{T}}-$ordered parton showers to model additional radiation in the leading-logarithmic approximation. The hadronisation model used is the Lund string model. MPIs are also simulated.

HERWIG **[79, 80]** is a general purpose Monte Carlo generator, which uses a LO $2 \to 2$ matrix element supplemented with angular-ordered parton showers in the leading-logarithm approximation. The cluster model is used for the hadronisation. The UE is modeled using an external package called JIMMY [81].

HERWIG++ **[82]** is based on the event generator HERWIG, but redesigned in the C++ programming language (HERWIG is programmed in Fortran). The generator contains a few modeling improvements. It also uses angular-ordered parton showers, but with an updated evolution variable and a better phase space treatment. The cluster model is also used for hadronisation. The UE are described using a multiple partonic interactions model [83].

ALPGEN **[84]** is a specialized tree matrix-element generator for hard multi-parton processes ($2 \to n$, with $n \leq 9$) in hadronic collisions. It is interfaced to HERWIG to produce

angular-ordered parton showers in leading-logarithmic approximation or PYTHIA to produce p_{T}−ordered parton showers. Parton showers are matched to the matrix element with the MLM matching scheme. The hadronisation process is simulated with HERWIG, using the cluster model. MPIs are modeled using JIMMY.

MC@NLO [85, 86] is a Fortran package which allows to match NLO QCD matrix elements consistently into a parton shower framework. In order to reproduce the NLO corrections fully, some of the configurations have negative weights. The shower and hadronization can be implemented using HERWIG or HERWIG++. The NLO expansion of the hard emissions needs to be evaluated for each showering program used.

MCFM [87] is a parton-level Monte Carlo generator. It is designed to calculate cross-sections for a wide range of processes at hadron-hadron colliders. For most processes, matrix elements are included at NLO and incorporate full spin correlations.

MADGRAPH [88] is a parton-level Monte Carlo generator. It generates, and calculates $2 \rightarrow n$ (with $n \leq 6$) matrix elements at LO for several processes. It is interfaced to PYTHIA to produce p_{T}−ordered parton showers.

ACERMC [89] is a parton-level Monte Carlo generator specifically used in the generation of SM background processes in pp collisions at the LHC. It provides a FORTRAN library of the matrix elements and phase space information for generation of a given process. It can be interfaced with either PYTHIA, HERWIG or a third Monte Carlo generator called ARIADNE (not described here) in order to simulated the initial and final state radiation, hadronisation and decays processes.

POWHEG [90] is a parton-level Monte Carlo generator. It allows to interface NLO calculations with a parton shower framework. It generates the hardest emission first, with NLO accuracy independently of the parton shower generator used. It can be interfaced with several parton shower generators as HERWIG, PYTHIA, etc.

3.3 The ATLAS detector simulation

The ATLAS collaboration has developed a detailed simulation of the ATLAS detector based on GEANT4 [91, 92] with an accurate description of the detector material and geometry. It propagates the generated particles through the ATLAS detector and simulates their interactions with the detector material. The energy deposited by particles in the active detector material is converted into detector signals with the same format as the ATLAS detector read-out. The simulated detector signals are in turn reconstructed with the same reconstruction software as used for the data. The tunable parameters in GEANT4 have been fixed according to the results from "test-beam" analyses. On the other hand, the accuracy of the detector geometry model and in the description of the basic properties of detector signals in the simulation have also been evaluated using 2010 data [93, 1, 94, 95, 96, 97].

The detector is described in terms of around 4.8 millions of volumes to match the real detector as close as possible. The detector simulation is based on two databases:

- The *geometry database* which contains information about the dimensions, positions and material properties of each detector volume.
- The *conditions database* which contains information about dead channels, temperature and misalignments in the real detector at a given time.

Using the conditions database, certain detector failures and beam conditions can be introduced in the simulation. Some of these conditions and failures vary from run to run. The simulation reflects the real experimental conditions as best known at the time of the Monte Carlo processing. For this reason, sometimes certain corrections are applied to the Monte Carlo simulation to reduce the differences with respect to the real data conditions. Monte Carlo data need to be produced at regular intervals, using information from the recent data taking to improve the description of the physics processes and of the detector geometry and conditions. In 2011, two different Monte Carlo and data reprocessings, also called releases, were done. In Chapter 6 the release 16 is used, which includes the first 2.05 fb^{-1} collected in 2011 and the simulation is denoted with the ATLAS tag **mc10b**. It has three simulated bunch trains with 225 ns separation between the trains. Each train has 36 filled bunches with 50 ns separation between bunches. These conditions were close to the real data taking collisions at the beginning of 2011. On the other hand, in Chapter 7 the release 17 is used. The later includes the whole dataset collected in 2011, 4.7 fb^{-1} and the Monte Carlo simulation is labelled with the ATLAS tag **mc11c**. In the mc11c simulation the number of proton bunch trains increased to four with respect to the mc10b simulation, to better model the collision environment observed at the end of 2011.

The output of the simulation is given in three different levels:

Parton level: refers to the partons at the matrix-element level which causes the particle shower due to the fragmentation process.

Particle or truth level: refers to the physics objects reconstructed after the fragmentation process and before any detector simulation.

Reconstructed level: refers to physics objects reconstructed from the detector signals.

3.4 Monte Carlo simulation weighting and corrections

After the detector simulation step, there are some corrections that need to be applied to the simulated samples to better match the real data. The most important correction to be done normalize each simulated sample to its theoretical cross section and to the number of expected data events. Usually, during the generation process a big number of events are produced to assure that different kinematic configurations of the process can be studied. The number of generated events does not correspond to the number of expected events in data. To get the right overall normalization an event weight, w, is applied to the Monte Carlo. It is given by:

$$w = \frac{\sigma \times k \times \mathtt{L}}{N}, \tag{3.7}$$

where σ is the cross section for the corresponding sample, $\mathtt{L}$ is the integrated luminosity, N is the number of generated events and k is known as the K-factor. The K-factor is a corrective factor to account for higher-order corrections. For example, they can be used to change a cross section from its LO value to its NLO order value. Negative weights associated with MC generators (as is the case for the MC@NLO) need to be taken into account in the event weight calculation. A second event weight applied correspond to the pile-up weight. It reweights the Monte Carlo to match the luminosity profile of the data, i.e. adjusts the number of simulated events to the expected number of events under the real pile-up data conditions. Extra event weights and corrections to the particle energies related with the reconstructed objects will be described in Chapter 4.

Chapter 4

ATLAS Event Reconstruction

Contents

The outputs of the digitization process of the detector signals generated by the particles produced in real/simulated pp collisions are processed by a serie of algorithms in order to built physics objects up. This step is not done in real time and thus is known as offline event reconstruction. The result is a set of physics objects with four-momenta that can be used directly in physics analyses. The reconstruction of tracks, primary vertices, electrons, muons, neutrinos and jets as well as the trigger chains used in the different analyses of this thesis will be described in the following sections. Taus can also be produced in the W boson decay from the top quark. They have a short lifetime and decay to quarks or to an electron or muon plus neutrinos. As there is no commissioned $\tau-$tagging algorithms in ATLAS yet, we do not specifically select and/or reconstruct taus in the analysis. A complete and detailed study of the expected performance of the ATLAS reconstruction is available at Ref. [98].

4.1 Data quality

The events where the relevant ATLAS subdetectors were not operational can not be used for physics analyses. In ATLAS, each subsystem is in charge of setting its own data quality and integrity flags for each LB. This information can be used to create a list of LB usable for analyses, called Good Runs List (GRL). Each analysis uses a GRL to reject those events affected by issues in the relevant subdetectors. In the studies presented in Chapter 5, the subdetectors of interest are the calorimeters. While in the studies presented in Chapters 6 and 7 nearly the full detector is relevant.

In 2011, an exception in the data quality assessment was the "LAr hole" issue. The information of 6 front end boards in the LAr calorimeter was lost due to a problem with their controller board. It created a hole in the detector data collection. It persisted for a few months before being fixed, affecting around 948.6 pb^{-1} of data. It represents 46% (20%) of data used in Chapter 6 (7). In release 16 of the ATLAS software, this issue was not modeled in the simulation. Instead of removing all the events affected, as would have been the procedure using a GRL, it was decided to remove only those events where the object reconstruction was affected by the issue. In release 17 of the ATLAS software, the issue was simulated in the Monte Carlo samples used. Therefore, no correction for the acceptance loss was needed.

4.2 Trigger chains

For the analyses presented in Chapters 6 and 7 it was decided to use single lepton triggers, since they are designed to trigger high-p_{T} events involving leptons, such as $t\bar{t}$ events. While in Chapter 5 level 1 minimum bias trigger or the level 1 and level 2 calorimeter triggers were used. The trigger requirement is applied to data and simulated samples. The trigger signatures used in this thesis varied for each data taking periods. Therefore, they will be described at the beginning of each analysis chapter.

4.3 Tracks

Tracks represent the trajectory of charged particles inside the detector. They are reconstructed using information from the ID (see Section 2.2.1). A precise track reconstruction is important to achieve a high vertex reconstruction efficiency and high precision in the particle momentum measurement. In ATLAS, tracks are parametrized by 5 parameters defined at the track's trajectory point closest (unless otherwise stated) to the center of the beam-spot: radial, d_0, and longitudinal, z_0, impact parameter, azimuthal ϕ and polar θ angle and charge signed inverse transverse momentum $\frac{q}{p_{\mathrm{T}}}$. The beam spot is the region where both beams interact, which does not correspond exactly to the geometrical center of the ATLAS detector. These parameters are represented in Figure 4.1. In general, the track reconstruction is done in three main steps [99]:

Track finding: assignment of ID hits to track candidates.

Track fit: determination of track parameters and their errors. Provides track's fit quality variables.

Test of track hypothesis: check the track candidate quality and the overlap with others tracks candidates.

In ATLAS the track finding and fit steps are merged. The hits from the pixel detector and the SCT are transformed into three-dimensional space points. Then, the *inside-out* algorithm [99] is applied. It consists in adding the three-dimensional space points one by one moving away from the interaction point to form a "road". It uses a combinatorial Kalman filter. The Kalman filter algorithm adds three-dimensional space points iteratively and fits simultaneously the track candidate. Cuts on the quality of the fit are applied to eliminate poor quality tracks and to avoid overlaps with others tracks candidates. The selected tracks are extended into the TRT and fitted again to get the final values of the track parameters. The mean energy loss in the detector material, the multiple scattering, the Bremsstrahlung effect and the changes in the magnetic field along the track trajectory are taken into account during the track fitting process. The inside-out algorithm provides the best reconstruction efficiency of primary charged particles directly produced in a pp collision or from decays or interaction of particles with a short lifetime ($< 3 \times 10^{11}$ s).

In order to better reconstruct secondary charged particles, produced in the interaction of primaries (with a lifetime $> 3 \times 10^{11}$ s), or conversion candidates an additional track finding algorithm, called *outside-in*, is applied [99]. The track finding process starts with TRT segments not used by the inside-out algorithm. They are then extended to the SCT and pixel detector. During 2011 data taking the detector occupancy increased significantly. Under these conditions the possibility of having incorrect hits assignments and more fake tracks from random hit combinations increase. The performance of the track reconstruction at ATLAS has been recently studied in the 2011 high pile-up environment [100]. The efficiency remains almost unchanged.

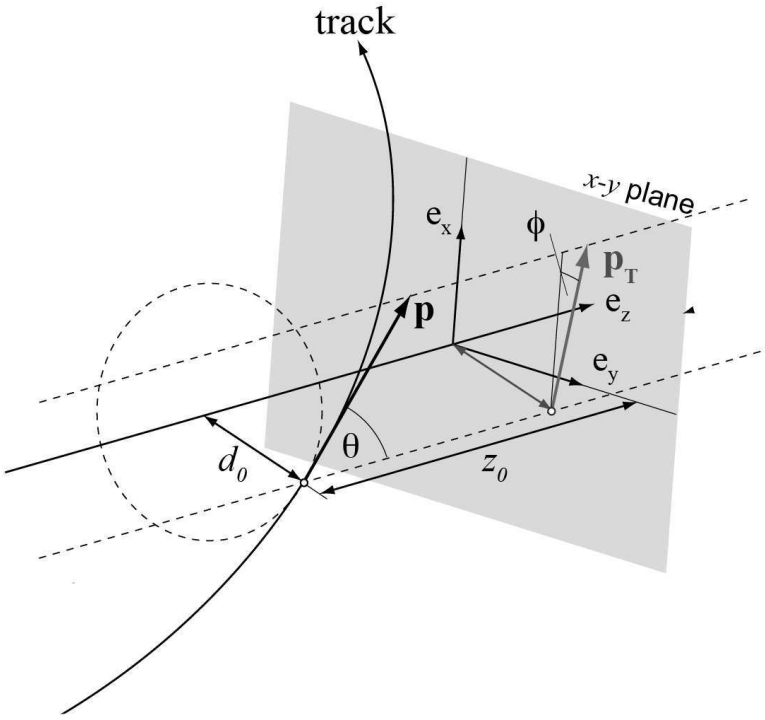


Figure 4.1: A track can be parametrized at one given point by its position, transverse momentum and charge. The position at any other point can be calculated if the magnetic field and the detector material is known.

However, the fraction of combinatorial fake tracks increases with the average number of pp collisions per bunch crossing at the time of the recorded event, $< \mu >$.

4.4 Primary vertices

There can be two different kinds of vertices in an event. First, the *primary vertices* (PV) which correspond to the collision point of beam particles. They are characterized by having many associated particles, thus a high track multiplicity. In an event there is one hard-scatter PV, while the rest are associated, by conventions, to pile-up interactions. Second, the *secondary vertices* correspond to decay of short-lived particles, which decay at a measurable distance from the PVs. The track multiplicity for secondary vertices is lower. See Section 4.8 for a more detailed description of the secondary vertices reconstruction process. In general, the PV reconstruction is done in three steps:

Vertex finding: assignment of reconstructed tracks to PV candidates.

Vertex fit: reconstruction of the PV position, calculation of its error matrix, estimation of the fit quality and optional re-fit of the associated tracks' parameters to constrain them to originate from the corresponding PV and not from the beam spot.

Test of vertex hypothesis: check the vertex candidate quality and the overlaps with others vertices candidates.

In ATLAS the PV reconstruction is done using an iterative vertex finding algorithm [101]. First, a vertex seed is found by looking for the maximum of the tracks' z_0 distribution. An

iterative χ^2 fit is used to fit the seed and the surrounding tracks. The matrix errors of the tracks are properly taken into account during the vertex fit. Tracks incompatible with the PV candidate (displaced by more than 7σ from the vertex) are used to seed a new PV. This procedure is repeated until no unassociated tracks are left or no additional vertex can be found. PVs are required to have at least two associated tracks. The same track can be associated to multiple vertices. The PV with the largest sum of squared tracks' transverse momenta $\sum p^2_{\mathrm{T,tracks}}$ is chosen as the hard-scatter PV. In high pile-up environment, the increasing number of fake tracks increases the probability to reconstruct a fake vertex. Furthermore, the common presence of nearby interactions increases the probability of reconstructing only one vertex out of several. Studies using 2011 data have shown that the PV efficiency reconstruction decreases with increasing $<\mu>$ [100]. Some quality criteria are applied to the tracks used in the PV reconstruction process. They vary from one analysis to the other. For the studies presented in Chapter 7, tracks were required to have at least 7 SCT+pixel detector hits, at most two holes in the pixel detector, $p_{\mathrm{T}}^{track} > 400$ GeV, $|d_0| < 2.5$ mm (w.r.t. PV) and $|z_0 \times \sin(\theta)| < 2$ mm (w.r.t. PV).

4.5 Electrons

4.5.1 Electron reconstruction

The distinguishing signature of an electron is a curved track in the inner detector and a narrow shower in the EM calorimeter. Electrons can be produced in the hard-scatter interaction, but also inside jets or in photon conversions. Then, there is a probability of misclassifying narrow jets or photons as electrons. During the electron reconstruction and identification several criteria are applied to determine if a given energy deposit and associated track were produced in the hard-scatter interaction or not. The electrons in the central region are reconstructed using an algorithm that combines the information from the EM calorimeter and the ID. Other algorithms reconstruct electrons in the forward region using only information from the EM calorimeter. In this thesis only electrons in the central region will be used. Therefore, the reconstruction of electrons in the forward region won't be discussed.

Electron reconstruction in ATLAS in the central region ($|\eta| < 2.47$) is done using the sliding-window algorithm [102]. It searches for clusters in the middle layer of the EM calorimeter with a total $E_{\mathrm{T}} > 2.5$ GeV. The window used to defined the clusters has a size of 3×5 in middle layer cell units ($\Delta\eta \times \Delta\phi = 0.025 \times 0.025$). If the seeded cluster is matched to a pair of tracks originating from a reconstructed photon conversion vertex, it is tagged as a converted photon. Otherwise, if the algorithm is able to match a track from the ID with the seeded cluster, it is tagged as an electron candidate. The matching is done in an $\Delta\eta \times \Delta\phi$ window of 0.05×0.10, to account for bremsstrahlung losses. The track momentum is required to be compatible with the cluster energy. In case that several tracks are matched to the EM cluster, the tracks with hits in the silicon detectors are preferred and the closest in ΔR is chosen. In addition, information from the TRT can be used to enhance the separation of electron candidates from pions. The final clusters are built around the seeded clusters matched with a track, by including all cells from different EM calorimeter layers located inside a rectangle centered on the seed position. The rectangle size depends on the position in the calorimeter of the seeded clusters (barrel or endcap). The energy of the cluster is calibrated to the EM energy scale, which was derived from MC based corrections to account for energy loss in passive material, test-beam measurements, and measurements of $Z \to ee$ decays for final calibration [103, 104].

4.5.2 Electron identification

Electrons can be distinguished from hadrons since EM showers deposit most of their energy in the second layer of the EM calorimeter. The width of electron showers is narrower than for hadrons. The ratio of the transverse energy reconstructed in the first layer of the hadronic calorimeter to the transverse energy reconstructed in the EM calorimeter, known as hadronic leakage, is smaller for electrons than for hadrons. Also the ratio of the energy reconstructed in the EM calorimeter to the track momentum E/p can be used as a discriminant variable, since it is smaller for charged hadrons. The most difficult task is to distinguish electrons from π^0s and ηs. They decay into two photons which form two close EM showers indistinguishable in the second EM calorimeter layer. In this case, the first layer of the EM calorimeter can be used due to its high granularity, to identify the two maximum in the π^0 or η shower corresponding to the two photons.

In ATLAS, there are 6 different series of cuts used in the electron identification process that provide good separation between electrons and jets faking electrons: *loose*, *loose++*, *medium*, *medium++*, *tight* and *tight++* [102]. In general, each one adds to the previous some additional requirements. The ++ menu was incorporated in release 17, in order to accomplish the trigger bandwidth restrictions for high luminosity. Its performance is better than the standard menu one. The discriminating variables used are defined using calorimeter and ID information.

Loose: is based on calorimeter information only. It requires electron candidates with $|\eta| < 2.47$ with low hadronic leakage and cuts on shower shape variables, derived from the energy deposits in the second layer of the EM calorimeter. The loose criteria provides a high identification efficiency. But the expected background rejection, i.e. the jet rejection, is low, about 500.

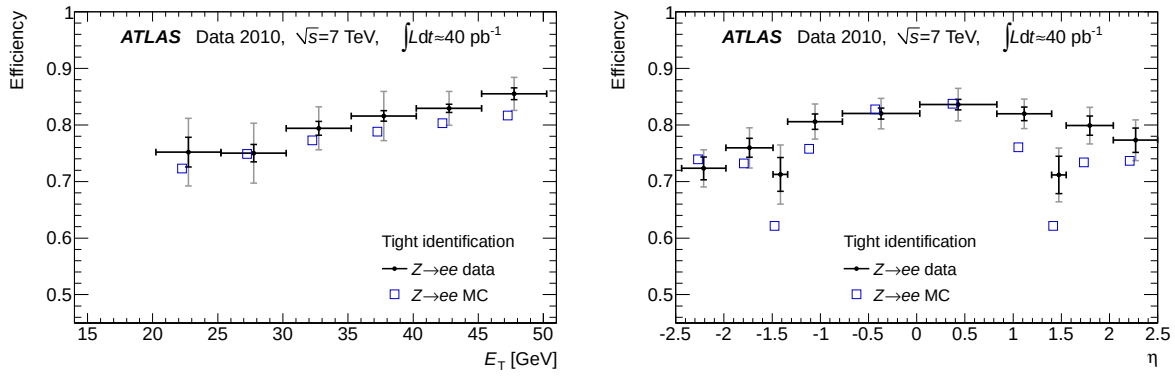
Loose++: adds additional cuts to the loose selection. It adds requirements related to the matched track: at least 1 hit in the pixel detector, at least 7 hits from both the pixel and SCT and the distance in $|\eta|$ between the cluster and the extrapolated track in the first EM layer has to be smaller than 0.015. Its efficiency, measured in $Z \rightarrow ee$ events, is close to the loose one (93%-95%) with a higher expected rejection of about 5000.

Medium: adds additional criteria related to the shower shape calculated using the first EM layer and to the deviation in the energies of the largest and second largest deposits in this layer, allowing discrimination against π^0s and ηs. In addition, the absolute value of the track's transverse impact parameter, $|d_0|$, is required to be lower than 5 mm and the distance in $|\eta|$ between the cluster and the extrapolated track in the first EM layer lower than 0.01. Its efficiency is of about 88% and has a rejection higher than the one achieved by the loose++ selection.

Medium++: requires at least one hit in the B-layer (first pixel detector layer) to reject electrons from photon conversions. Tracks having a low fraction of high-threshold TRT hits are rejected to decrease the contamination from charged hadrons. $|\Delta\eta|$ between the cluster and extrapolated track in the first EM layer is lowered to 0.005. Tighter shower shapes cuts for $|\eta| > 2.01$ are used. It has an efficiency of around 85%, with a expected rejection closer to 50000.

Tight: requires that $|\Delta\phi|$ and $|\Delta\eta|$ between the cluster and the matched track has to be less than 0.02 and 0.005, respectively. A requirement on E/p is introduced. The $|d_0|$ requirement is tightened (to be less than 1 mm), as well as the fraction of high-threshold TRT hits. The identification efficiency is around 75% (see Figure 4.2), with a rejection higher than the medium++ one.

Tight++: only adds asymmetric $\Delta\phi$ track-cluster matching cuts. It has an efficiency slightly better than the one for tight selection and a slightly better rejection too.



(a) (b)

Figure 4.2: Tight electron identification efficiencies measured from $Z \to ee$ events in data and MC as a function (a) of E_{T}, integrated over $|\eta| < 2.47$ excluding the transition region $1.37 < |\eta| < 1.52$ and (b) of eta and integrated over $E_{\mathrm{T}} > 20$ GeV [102].

The performance of electron reconstruction, trigger and identification was evaluated in 2010 data and MC using $Z \to ee$ and $W \to e\nu$ events [102]. To suppress the background due to non-prompt leptons, e.g. from decays of hadrons (including heavy flavour) produced in jets, the leptons in the event are usually required to be isolated. A calorimeter isolation, a track isolation or both can be applied. The calorimeter isolation is estimated using the energy in a cone of $R = 0.2$ centered around the electron after the subtraction of the energy associated with the electron itself, `EtCone20`. Track isolation is calculated using the scalar sum of tracks p_{T} in a cone of $R = 0.3$ centered around the electron without including the electron p_{T} itself, `PtCone30`. The calorimeter isolation variables usually include a correction for the increase in the electron's energy in the isolation cone with electron p_{T} (transverse shower leakage) and for additional energy deposits from pile-up events.

4.5.3 Electron scale factors and energy corrections

In release 16, electron identification scale factors were calculated only as a function of η_{cluster}. The differences in trigger and reconstruction efficiency were taken into account using flat scale factors of 0.995 ± 0.01 and 1.013 ± 0.015, respectively [102]. They have been estimated using tag & probe methods in $Z \to ee$ and $W \to e\nu$ samples. In release 17, trigger, reconstruction and identification (including isolation) efficiency scale factors for electrons (with its respective systematic uncertainties) were derived as a function of η_{cluster} and E_{T} [5]. They are applied to simulation as an event weight to take into account the differences in lepton efficiencies between data and simulation. These scale factors are around 1 and their impact in the analysis is small.

The smearing or scaling of the reconstructed objects at the analysis level is a common procedure. They are used to match the object energy in simulation to the one in data, to match the object energy to a known quantity or to implement an uncertainty in the analysis. The smearing process consists in changing the object energy distribution using random numbers from a given distribution, usually a gaussian. In the case of electrons, the EM electron cluster energy in data was corrected by applying energy scales as a function of η, ϕ and E_{T} to match the Z boson peak mass. They were obtained from $Z \to ee$, $J/\Psi \to ee$ or E/p studies using isolated electrons

from $W \to e\nu$. The cluster energy was smeared in Monte Carlo samples to match the energy resolution in data and to adjust the width of the Z peak.

4.6 Muons

4.6.1 Muon reconstruction and identification

When muons pass through the detector they leave curved tracks in the ID and in the MS. They traverse the detector as minimum ionizing particles and therefore can deposit a small fraction of their energy in the calorimeters. This information can be used to improve the energy resolution measurement. Similar to electrons, the muons can be identified in different categories. *Standalone muons* are reconstructed using MS tracks backtracked to the interaction point. *Combined muons* are obtained by matching standalone muons with ID tracks and refitting the combination. *Tagged muons* are built by extrapolating ID tracks to the MS and combining them with segments reconstructed in MS stations. Finally, *calorimeter muons* are reconstructed extrapolating ID tracks to the calorimeters and combining them with calorimeter energy deposits.

In this thesis only tight combined muons are used. They are reconstructed using the Muid algorithm [105]. The track finding process in the MS starts by searching for straight track segments in the RPC and TGC. The search is done in regions of size $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$ where some trigger activity has been detected. A track is formed with two or more segments in different muon chambers using a least-square fitting method. Tracks candidates are then extrapolated to the beam spot. The momentum is corrected for the energy loss in the calorimeters and the magnetic field. The ID muon tracks are reconstructed separately, as described in Section 4.3. A χ^2 test is performed on the ID and MS tracks to form a final muon track. It is defined from the difference between the respective track extrapolated coordinates weighted by their combined covariance matrices. If a combined track cannot be formed, the particle is not considered to be a muon. If several combinations are possible, the ID/MC track pair with the best χ^2 is chosen as the muon track candidate. Some quality cuts are applied to the ID track candidates:

- At least one hit in the B-layer of the pixel detector.
- At least two pixel hits.
- The number of pixel hits plus the number of crossed dead pixel sensors ≥ 2.
- The number of SCT hits plus the number of crossed dead SCT sensors ≥ 6.
- The number of pixel holes plus the number of SCT holes ≤ 2. A hole is the region where the module did not respond as expected, eventhough the rest of the modules elsewhere along the track did.
- There is a complex requirement on the number of TRT outliers (m) divided by the number of TRT outliers plus the number of TRT hits (n) that depend on η. Outliers are hits that deviate from the track. For $\eta < 1.9$ is required $n > 5$ and m/n smaller than 0.9. For $\eta \geq 1.9$, if $n > 5$ then m/n is required to be < 0.9.

The coverage is limited by the ID acceptance to $|\eta| < 2.5$. The momentum of the combined muon is then calculated by the weighted average of the ID and the MS momentum measurements, which improve the resolution with respect to the one obtained for standalone muons at low p_{T}. The performance of muon reconstruction, trigger and identification was evaluated in 2010 [105] and in 2011 (see Figure 4.3) data and MC using $Z \to \mu\mu$ events. A calorimeter and a track isolation can be defined for muons in the same way as it was done for electrons.

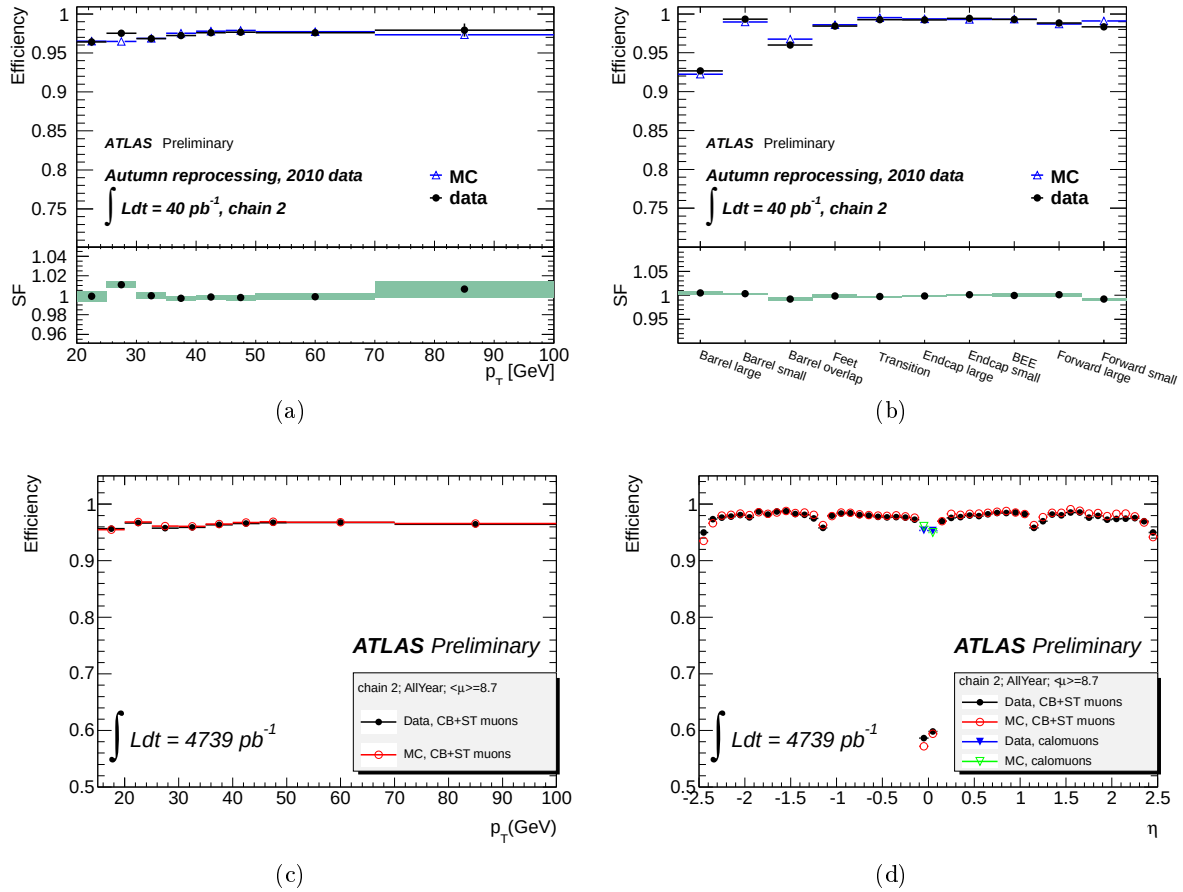


Figure 4.3: Combined Muid efficiencies measured from $Z \rightarrow \mu\mu$ events in data and MC as a function (a, c) of E_{T} and (b, d) of η, using 2010 data (up) [105] and 2011 data (bottom).

4.6.2 Muon scale factors and energy corrections

Trigger, reconstruction and identification (including isolation) efficiency scale factors derived using $Z \rightarrow \mu\mu$ events, are applied to Monte Carlo as a function of η, ϕ and the data period. Energy scale shifting and resolution smearing to correct the reconstructed muon momentum in MC are applied to match the energy scale and resolution measured in data $Z \rightarrow \mu\mu$ events. As tracks are build from the ID and MS, there are separate corrections on the tracks in each region.

4.7 Jets

At high energy pp collisions the presence of partons is overwhelming. Due to colour confinement the partons hadronize. While the resulting bunch of particles passes through the ATLAS detector, they produce tracks in the ID and energy deposits inside the calorimeters. These detector signals allow the reconstruction of track jets (reconstructed using track information) and calorimeter jets (reconstructed using calorimeter information). The different analyses presented in this thesis make use of calorimeter jets. This section will then focus in explaining the jet reconstruction process for calorimeter jets only. It consists in three steps: the definition of calorimeter signals, the use of a jet reconstruction algorithm to group the calorimeter signals and finally the jet calibration which corrects the jet energy and momentum for the effects of ATLAS calorimeters non-compensation, dead material, leakage, out of cone and other thresholds effects.

4.7.1 Jet reconstruction

4.7.1.1 Inputs to calorimeter jet reconstruction

In a first step, calorimeter cells are combined and the resulting clusters will be used as inputs for the jet reconstruction algorithm. ATLAS provides two different cells clustering algorithms and therefore two types of inputs to jet reconstruction:

Topological clusters or topoclusters are a dynamically formed combination of cells around seed cells that exceed a given signal-to-noise ratio threshold. The seeds are defined to be the cells with $|E_{cell}/\sigma_{cell}^{noise}| > 4$, where E_{cell} is the cell energy and σ_{cell}^{noise} is the RMS of the cell noise distribution. Subsequently, their neighboring cells are included if their signal-to-noise ratio exceeds a second threshold $|E_{cell}/\sigma_{cell}^{noise}| > 2$. Finally all cells neighboring the formed topoclusters are added to the topocluster. Topoclusters are defined as massless. Their energy is obtained summing up the energy of all the cells included. Their direction is calculated from weighted averages of the pseudorapidity and azimuthal angles of the constituent cells relative to the nominal ATLAS coordinate system. The weight used is the absolute cell energy. Because of calorimeter noise fluctuations cluster can have a negative energy. Negative energy clusters are rejected entirely from the jet reconstruction since they do not have physical meaning.

Noise Suppressed Towers are constructed by projecting calorimeter cells onto a grid with tower bin size $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. When cells larger than 0.1 exist, like in the third layer of the tile calorimeter, they are splitted between towers, and so is their energy, in a proportional manner. The towers are built using only cells belonging to topological clusters. Therefore, the same noise suppression is used in both cases. Towers are also defined to be massless and their energy and direction are calculated in the same way as for topoclusters. Negative energy towers are rejected entirely from the jet reconstruction. The difference between towers and clusters is illustrated in Figure 4.4. Jets built of towers are not used in this thesis.

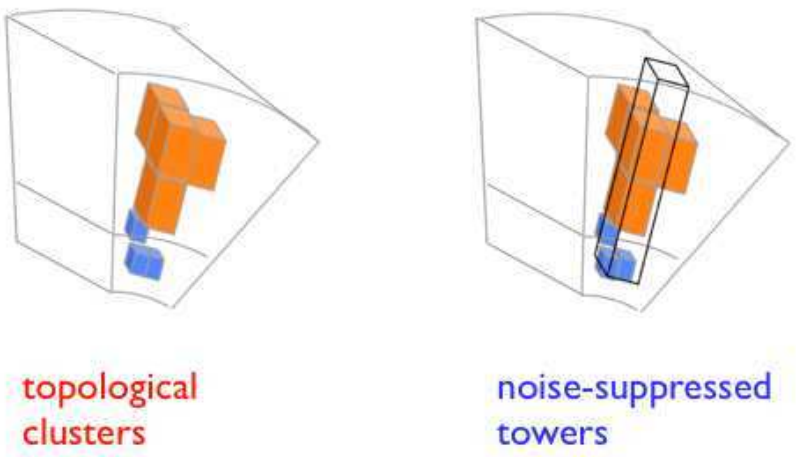


Figure 4.4: Illustration of how cells may build a topological cluster (left) and how noise suppressed towers are built from those cells (right).

4.7.1.2 Jet reconstruction algorithms

Jet reconstruction algorithms allow to associate the energy deposits in the calorimeters to a jet. A good jet algorithm should give a stable and precise description of QCD interactions during the pp collision and therefore has to fulfill certain conditions:

- Collinear safety, which means that the splitting of one particle into two collinear particles has no effect on the reconstruction.
- Infrared safety, which means that the presence of additional soft particles between jet components does not affect the jet reconstruction.
- Effects of resolution and other detector effects (e.g. noise) should affect the jet reconstruction as little as possible.
- Invariance under Lorentz boosts along z coordinate.
- Minimum computer resources used.

The jet reconstruction algorithm used in this thesis is called the anti$-k_t$ algorithm [106], which is a sequential recombination algorithm. Sequential recombination algorithms take topoclusters or towers as input and combines them to form jets according to a distance parameter defined below. For all inputs i, and pairs ij two different distances are defined:

$$d_{ij} = min(p_{\mathrm{T},i}^{2p}, p_{\mathrm{T},j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \tag{4.1}$$

$$d_i = p_{\mathrm{T},i}^{2p}, \tag{4.2}$$

where p_{T} is the transverse momentum of the input i, $\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ is the distance between a pair of inputs in the $y - \phi$ space, R and p are parameters of the algorithm. d_{ij} represent the distance between a pair of inputs i and j, while d_i the distance between the input i and the beam axis in the momentum space. The algorithm calculates $min(d_i, d_{ij})$. If $min(d_i, d_{ij}) = d_i$, the input i is said to form a jet and is removed from the list of inputs. If $min(d_i, d_{ij}) = d_{ij}$, the inputs i and j are combined into one single input using the E-scheme (sum of four-momentum of each input). The combined input is put into the list of possible inputs, while i and j are removed. The algorithm proceeds until no inputs are left, which means that all inputs in the event will end in a jet.

The parameter p defines the kind of algorithm:

- $p = 1$: k_t algorithm [107]
- $p = 0$: Cambridge/Aachen algorithm [108]
- $p = -1$: anti$-k_t$ algorithm [106]

while R characterizes the size of the jet in the $y - \phi$ space. The anti$-k_T$ algorithm works in the inverse transverse momentum space and has three main advantages. First, it clusters nearby particles, ensuring infrared safety. Second, soft inputs prefer to cluster with hard inputs instead of clustering with other soft particles. Third, the anti$-k_T$ algorithm is seedless and all hard inputs within $\Delta R_{ij} < R$ will be combined into one jet, ensuring the collinear safety. In the studies presented in the following chapters, two resolution parameters are used: $R = 0.4, 0.6$.

The topoclusters and towers are defined massless. However, their distribution inside the reconstructed jet leads the jet to have a given mass. In Monte Carlo simulations track jets and

calorimeter jets are reconstructed as in real data. In addition, two other kinds of jets can be defined in simulation:

Parton jet refers to the parton at the matrix-element level which causes the particle shower due to the fragmentation process.

Particle or truth jet which is reconstructed from stable particles[1] produced by the hadronization models of the different Monte Carlo generators, before any detector simulation. This mainly includes electrons, photons, pions, kaons, protons and neutrons and their antiparticles. Neutrinos and muons are not included, since they do not leave any significant signal in the calorimeter.

In Figure 4.5 a schematic view of the different types of jets is shown. The jet reconstruction efficiencies were determined from data with a tag&probe method, using track jets, where the efficiency was defined as the fraction of probe track jets matching a corresponding calorimeter jet [109]. The difference between data and simulation is found to be small and within the uncertainties. Therefore, no scale factors were needed.

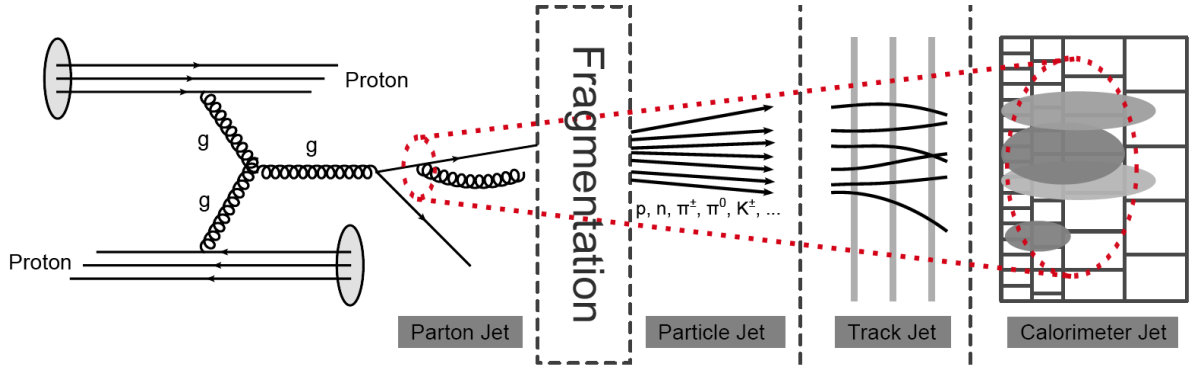


Figure 4.5: Schematic view of the different types of jets. Colorful particles, produced at the interaction point, create a bunch of colorless particles due to fragmentation. These particles will produce detector signals in the Inner Detector and Calorimeter, which can be reconstructed as track and calorimeter jets.

4.7.2 Jet calibration

The energy of the reconstructed jets does not correspond to the initial energy carried by the particles. Reconstructed jets need therefore to be calibrated to the correct energy scale. In general, the reference scale in the jet calibration process is given by the truth jets. Jets are initially reconstructed at the electromagnetic (EM) scale, which is the basic calorimeter signal scale for the ATLAS calorimeters. This means that the calorimeter signals are calibrated to properly reproduce the energy lost in the calorimeter by an electron, if the energy deposit came from an electron. The EM scale was obtained using test-beam measurements for electrons in the barrel [110, 103, 104] and the endcap calorimeters [111]. It has been validated using muons from test-beams and in cosmic-rays. The energy scale of the electromagnetic calorimeters has been corrected using the invariant mass of $Z \rightarrow ee$ events [102]. This EM scale calibration provides a very good description for energy deposits produced by electrons and photons, but not for deposits

[1] Stable particles refer to particles with a lifetime greater than 10 ps.

from hadronic particles like protons, neutrons, pions or kaons. This difference results from the fact that the EM scale calibration does not account for various detector effects:

- **Calorimeters non-compensation:** partial measurement of the energy deposited by hadrons.
- **Dead material:** energy losses in inactive or non instrumented regions of the detector.
- **Leakage:** energy deposits from particles which are not totally contained in the calorimeters.
- **Out of cone:** loss of energy deposits from particles inside the particle jet that are not included in the reconstructed jet.
- **Thresholds effects:** signal losses due to inefficiencies in calorimeter clustering and jet reconstruction.

ATLAS has developed several calibration schemes with different levels of complexity and different sensitivity to systematic effects [3], which made them complementary in the way they contribute to the understanding of the jet energy scale measurement:

EM+JES calibration is a Monte Carlo-derived jet calibration scheme used in ATLAS for the first analyses. EM+JES applies a simple jet-by-jet correction that restores the reconstructed jet energy to the particle jet energy. In this scheme each jet at the EM scale is scaled by a correction factor which is a function of the reconstructed jet energy and η. In addition to this energy correction a pile-up and a jet origin correction are also applied. See Section 4.7.4.1 for a more precise description.

Global Sequential (GS) Calibration is a Monte Carlo-derived jet calibration, which uses longitudinal and transverse properties of the jet structure sequentially to improve the resolution, while leaving the jet energy scale unchanged. In this scheme jets are found from clusters or towers, then the EM+JES calibration is applied and finally they are scaled by a jet-by-jet correction factor which depends on the jet p_T, η and several longitudinal and transverse jet properties. See Chapter 5 for a more detailed description.

Global Cell Energy-Density Weighting Calibration (GCW) [93, 1] attempts to compensate for the different calorimeter response to hadrons and electromagnetic particles by weighting each jet constituent cell. The weights, which depend on the cell energy density and the calorimeter layer only, are determined by minimizing the energy fluctuations between the reconstructed and particle jets in Monte Carlo simulation. Jets are found from topoclusters or towers at the EM scale, then cells are weighted and a final jet energy scale correction is applied to ensure that good linearity response is achieved.

Local Cluster Weighting Calibration (LCW) [93, 1] uses properties of topoclusters (such as their energy, depth in the calorimeter, cell energy density, fractional energy deposited in the calorimeter layer and energy measured around it) to calibrate them individually before applying jet reconstruction. These weights are determined from Monte Carlo simulations of charged and neutral pions. Similarly to the GCW scheme, a final correction of the jet energy is applied.

The EM+JES calibration is used for first physics analysis, due to its simplicity. The others calibration schemes are presently commissioned by ATLAS. The corrections applied by each

calibration schemes as well as the inputs used have been validated using data from pp collision at a center-of-mass energy of $\sqrt{s} = 7$ TeV [3]. Studies regarding their performance are shown in references [3, 112].

4.7.3 Jet selection

Jet quality criteria were applied to identify the so called *bad jets*. Some sources of bad jets are for example hardware problems, calorimeter showers induced by cosmic rays and beam remnants. The used criteria varied for each analysis presented in this thesis. Therefore, they will be described at the beginning of each analysis.

4.7.4 EM+JES calibration

4.7.4.1 Description of the calibration

The EM+JES calibration [3] consists of four consecutives steps:

Offset correction: jets built from topoclusters or noise suppressed towers at the EM scale are corrected by the offset correction to subtract the additional jet energy due to multiple pp interactions. This additional energy can come from multiple pp collisions within the same bunch crossing (this is known as *in-time* pile-up) or from pp interactions in a previous bunch crossing, if the previous bunch crossing happened within the read-out window of the calorimeters (known as *out-of-time* pile-up). For 2010 data only an in-time pile-up correction was derived as the effect of out-of-time pile-up was small, due to the "long" spacing between bunches [113]. For early 2011 data a pile-up correction was not included since the jet energy scale correction was calculated using Monte Carlo samples simulated with pile-up ($< \mu >= 8$ and a 75 ns of bunch spacing). For later 2011 data the offset correction was again introduced since the pile-up conditions increased. It was updated to account for bigger out-of-time pile-up effects too [114].

Jet Origin Correction: the jet direction is recalculated to point to the direction of the primary vertex with the highest sum of track transverse momentum in the event and not the geometrical center of the ATLAS detector. This is a valid assumption as long as the jet originates from that vertex, and not from pile-up. This correction improves the angular resolution while the jet energy is unaffected.

Jet Energy Scale Correction: the resulting jet energy is corrected using a $p_{\mathrm{T}}-$ and $\eta-$dependent correction to that of the corresponding truth jet. It is derived in Monte Carlo simulation. The correction is calculated using all isolated[2] calorimeter jets that have a matching isolated truth jet within $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$ (this value results in a reconstructed-to-truth jet match more than 99% of the time for jets with $p_{\mathrm{T}}^{\mathrm{truth}} > 20$ GeV). The jet energy response $\mathcal{R}$ is measured in bins of E_{truth} (energy of the matched truth jet) and η:

$$\mathcal{R} = E_{\mathrm{reco}}^{\mathrm{EM}} / E_{\mathrm{truth}}, \tag{4.3}$$

where $E_{\mathrm{reco}}^{\mathrm{EM}}$ is the energy of the calorimeter jet at the EM scale. $E_{\mathrm{reco}}^{\mathrm{EM}}$ is used directly to calculate the response because the Monte Carlo simulation used does not include additional pp interactions, so no offset correction is needed. The original detector η is used rather

[2] An isolated calorimeter (truth) jet is defined as a jet that has no other calorimeter (truth) jet with EM scale (truth) $p_{\mathrm{T}} > 7$ GeV within $\Delta R = 2.5R$, where R is the distance parameter of the jet algorithm

than the origin-corrected one, since it corresponds directly to a region of the calorimeter. The average jet response $\langle \mathcal{R} \rangle$ is defined as the mean of a gaussian fit to the peak of the response distribution for each $(E_{\text{truth}}, \eta)-$bin. In each $(E_{\text{truth}}, \eta)-$bin, the corresponding average calorimeter jet energy, $\langle E^{\text{EM}}_{\text{reco}} \rangle$ is calculated as follows:

$$\langle E^{\text{EM}}_{\text{reco}} \rangle = E_{\text{truth}} \times \langle \mathcal{R} \rangle, \tag{4.4}$$

so the response can be expressed as a function of $E^{\text{EM}}_{\text{reco}}$. Initially the response is derived in E_{truth} bins, since the the jet response distribution is only gaussian only in bins of E_{truth}. In bins of $E^{\text{EM}}_{\text{reco}}$, the jet response is distorted (not gaussian) due to the impact of the underlying p_{T} spectrum [115].

The calibration constants that relate the measured calorimeter jet energy to the truth jet energy $C(E^{\text{EM}}_{\text{reco}}, \eta)$ are defined as the inverse of the response:

$$C(E^{\text{EM}}_{\text{reco}}, \eta) = \langle \mathcal{R} \rangle^{-1}. \tag{4.5}$$

They are derived in 0.1-wide $|\eta|$ bins to capture the changing calorimeter geometry. Figure 4.6 shows the average jet energy response before calibration for jets in the Monte Carlo simulation as a function of $|\eta|$ for jets built of topoclusters with the anti-k_t algorithm with $R = 0.6$. The jet energy response $\eta-$dependence is due mainly to poorly instrumented regions of the calorimeters (especially the transition region between the hadronic endcap and the forward calorimeter).

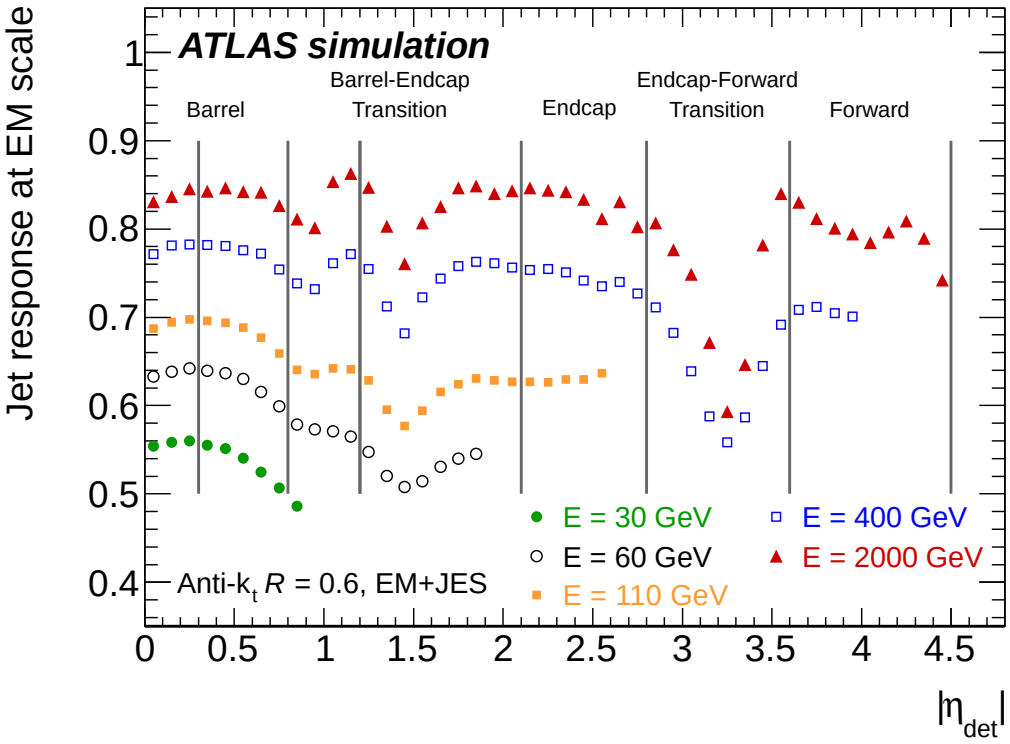


Figure 4.6: Average simulated jet energy response at the EM scale as a function of the EM+JES calibrated jet energy $E^{\text{EM+JES}}_{\text{reco}}$ (called E in the plot) and the detector pseudo-rapidity η_{det} [3].

Jet Position Correction: after the jet origin and energy corrections, a small η-dependent correction needs to be applied since the η measurement is affected by the calorimeter response η-dependence. When topoclusters and towers are clustered to form a jet, their energies are used as a weight in the calculation of the jet direction. Cells are reconstructed with a lower energy in non instrumented regions with respect to better instrumented regions (see Figure 4.6). Therefore, the jet direction will be biased towards better instrumented regions. This bias is corrected using a technique similar to the one used to correct the jet energy. The correction is derived as the average $\Delta\eta = \eta_{\text{truth}} - \eta_{\text{origin}}$ in (E_{truth}, η)-bins and is parametrized as a function of $E^{\text{EM}}_{\text{reco}}$ and the detector η. It is a very small (< 0.01) correction for most of the regions of the calorimeter, except in the transition regions.

Residual data calibration: in release 17 a new step was added to the EM+JES calibration chain. The residual data calibration is derived from in-situ measurements, where well calibrated objects are used as reference. Direct p_{T} balance techniques are used to adjust the jet calibration in the data by comparing the $p^{\text{jet}}_{\text{T}}/p^{\text{ref}}_{\text{T}}$ ratio in the data with to ratio in the simulation.

A detailed description of the associated jet energy scale systematic estimated in release 16 and 17 of the ATLAS software can be found in Ref. [116] and [3, 5], respectively.

4.7.4.2 Jet energy resolution

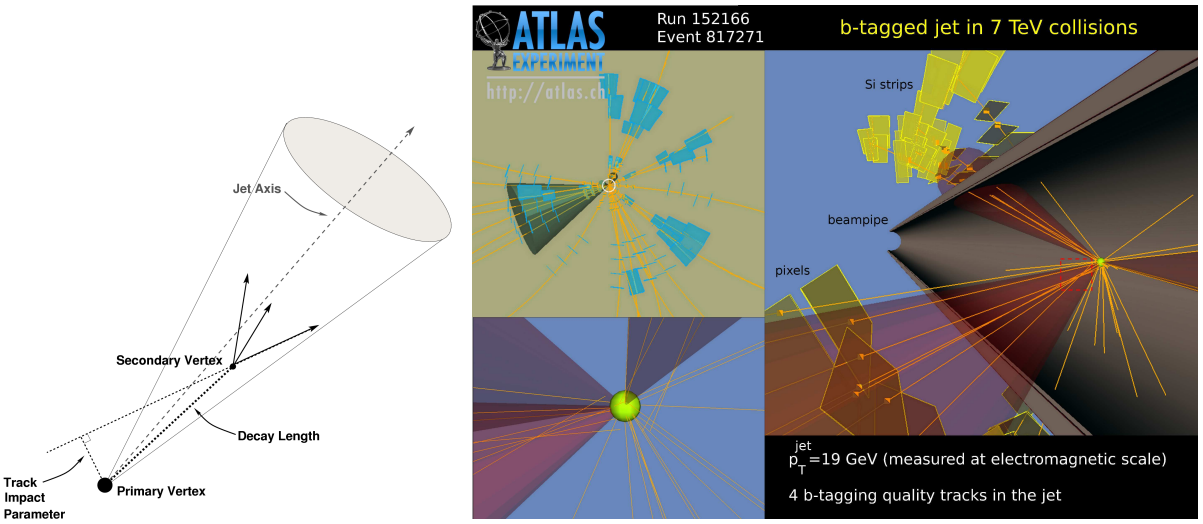
The jet energy resolution (JER) of jets at the EM+JES scale was measured with the di-jet balance and the bi-sector techniques [112, 5]. In release 17 the measured JER in data and simulation was found to agree within the uncertainties. Therefore, no smearing was applied to jets in simulation. Unfortunately, this was not the case in release 16. The energy of simulated jets was smeared according to a gaussian distribution in order to match the jet energy resolution in data.

4.8 $b-$tagging

4.8.1 $b-$tagging algorithms

The B hadron formed by the bottom quark has a relatively long lifetime of about 1×10^{-12} s and can travel around 3 mm before decaying. The identification of $b-$jets is very important for the discrimination of top quark analysis backgrounds with only light jets in the final state. $b-$tagging algorithms exploit the fact that a certain number of tracks point to a secondary vertex instead of pointing to the reconstructed primary vertices as shown in Figure 4.7 and that impact parameters of these tracks are large. The $b-$tagging procedure relies on the ID track reconstruction. Therefore it can be only applied to jets with $|\eta| < 2.5$. ID tracks are required to pass some quality criteria that depend on the different $b-$tagging algorithms. There are basically three kinds of $b-$tagging algorithms [117]:

Impact parameter-based algorithms: use the significance of the impact parameters, z_0/σ_{z_0} and d_0/σ_{d_0}, of each track contained in the respective jet to determine a likelihood corresponding to the $b-$jet tag probability. σ_{d_0} and σ_{z_0} correspond to the d_0 and z_0 errors, respectively. The algorithms `IP1D`, `IP2D` and `IP3D` use z_0/σ_{z_0}, d_0/σ_{d_0} and a combination of them, respectively. A jet is considered to be $b-$tagged if its calculated tag probability is above of certain threshold, known as the operating point of the $b-$tagging algorithm.



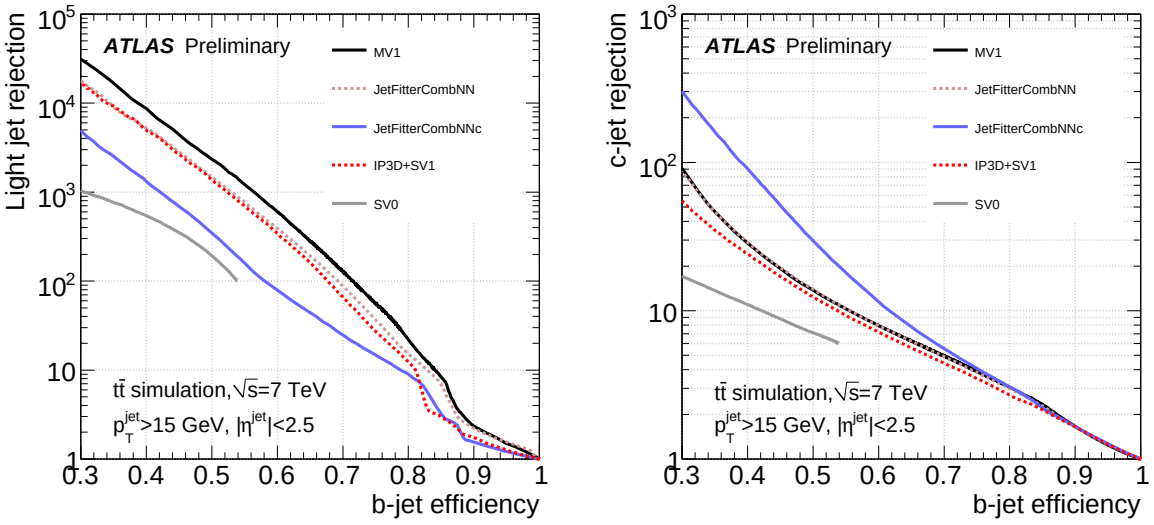
(a) (b)

Figure 4.7: (a) Schematic and (b) real collision view of a $b-$jet candidate. The event display shows a $b-$jet candidate reconstructed with the anti$-k_T$ algorithm, where the primary vertex is shown in the yellow circle and the secondary vertex in the dashed red box.

Secondary vertex-based algorithms: reconstruct the secondary vertex formed by the decay products of the $b-$hadron to increase the discrimination between $b-$jets and light jets. Secondary vertex related discriminant variables are used by the algorithms to calculate the $b-$tagging probability. The `SV0` algorithm uses the decay length significance, L/σ_L, measured in 3D and signed with respect to the jet direction. The `SV1` algorithm uses the number of tracks pairs contained in the secondary vertex (which is larger in the case of $b-$jets), the invariant mass of all the tracks associated to the secondary vertex (which is higher for $b-$jets than for light jets) and the ratio of the sum of the energies of the tracks in the vertex to the sum of the energies of all tracks in the jet ($b-$jet tracks contains in general a big fraction of the jet energy). Each one of these discriminating variables is converted into a likelihood ratio. The jet $b-$tagging weight is given by the sum of the logarithms of the individual weights.

Decay chain reconstruction: this is the case of the `JetFitter` algorithm. It uses a Kalman filter to fit the decay chain of the b and c hadrons, determining a common path between the primary vertex and the b and c hadrons vertices (inside the $b-$jet candidate). A likelihood is built to distinguish between $b-$ and light jets using the flight length significances of the vertices and the variables used by the `SV1` algorithm.

A high $b-$tagging efficiency associated to a high light jet rejection can be obtained when combining the algorithms described above. The `JetFitterCombNN` algorithm determines the jet $b-$tag weight according to a Neural Network combination of the weights from the `IP3D` and `JetFitter` algorithms. The `MV1` algorithm follows a similar procedure, but using the weights from the `JetFitter`, `IP3D` and `SV1` algorithms. Figure 4.8 shows the light jet and $c-$jet rejection power versus the $b-$tagging efficiency for different $b-$tagging algorithms.



(a) (b)

Figure 4.8: The (a) light jets and (b) $c-$jets rejection power versus the $b-$tagging efficiency for different $b-$tagging algorithms measured in simulated $t\bar{t}$ events [118].

4.8.2 $b-$tagging scale factors

The performance of the $b-$tagging algorithm is given by the efficiency with which a $b-$jets is tagged as such, by the $c-$tag efficiency, which is the equivalent quantity for jets originating from $c-$quarks, the $\tau-$tag, the light jet tag efficiency and by the corresponding mistag rates. They are not the same in data and simulation, so we need to apply scale factors to correct the simulation [119, 118, 120, 121]. In release 16, for tau jets the light jets efficiency and mistag SFs were used, but with twice the uncertainty, since the tau jets have few tracks and are more similar to light jets. For release 17, additional SFs for $\tau-$leptons were derived. $b-$tagging scale factors have values close to 1, but their systematic uncertainties are large.

4.9 Missing transverse energy and neutrinos

Neutrinos hardly interact with the detector material. Therefore, their presence can only be inferred through the total transverse momentum balance in the event. Since the hard interaction occurs between partons carrying only a fraction of the hadron's energy, the longitudinal component of the initial state energy is unknown. Only the transverse component of energy of the initial state is known to be zero. A quantity known as the transverse missing energy $E_{\mathrm{T}}^{\mathrm{miss}}$, defined as the energy imbalance in the transverse plane, can be defined. In ideal conditions, $E_{\mathrm{T}}^{\mathrm{miss}}$ should be zero when non interacting particles are not produced in the event. But when a non interacting particle is produced, it should point to the non interacting particle direction in the transverse plane with a magnitude close to its p_{T}. If there is more than one neutrino produced, there will be still one $E_{\mathrm{T}}^{\mathrm{miss}}$ but the information of the neutrinos' directions will be lost. There are many sources of fake $E_{\mathrm{T}}^{\mathrm{miss}}$ that need to be taken into account during its calculation: detector coverage, presence of dead, non instrumented or noisy regions and finite detector resolution.

There are two main algorithms available to reconstruct the $E_{\mathrm{T}}^{\mathrm{miss}}$. The first one uses topological clusters as inputs, while the second one uses reconstructed objects [122]. In this thesis, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is estimated from the sum of the topological clusters associated to the reconstructed

objects (high and low p_{T} jets, electrons), the energy from reconstructed muons in the x and y directions and clusters calibrated at the EM scale not associated to any object, called cell-out. Calorimeter cells in the topoclusters are associated with objects in a chosen order: electrons, jets and muons, to avoid double counting of the energy in the cells. Calorimeter cells are calibrated according to the reconstructed physics object to which they are associated. Cells belonging to jets with $p_{\mathrm{T}} > 20$ GeV are corrected to the EM+JES scale, while those belonging to jets with 7 GeV $< p_{\mathrm{T}} < 20$ GeV were included at the EM scale. Cells belonging to electrons with $p_{\mathrm{T}} > 10$ GeV are considered and calibrated with the default electron calibration (see Section 4.5). Calorimeter cells associated to the selected muons are subtracted. This definition is referred to as refined $E_{\mathrm{T}}^{\mathrm{miss}}$, since each topological cluster have been calibrated and corrected according to the reconstructed object to which they are associated. Another possibility is to use the topological clusters at the EM scale. When all of these contributions are summed, the result is expected to be zero. Any deviation from zero correspond to $E_{\mathrm{T}}^{\mathrm{miss}}$.

$$E_{x,y}^{\mathrm{miss}} = E_{x,y}^{\mathrm{electrons}} + E_{x,y}^{\mathrm{photons}} + E_{x,y}^{\mathrm{taus}} + E_{x,y}^{\mathrm{jets}} + E_{x,y}^{\mathrm{softjets}} + E_{x,y}^{\mathrm{muons}} + E_{x,y}^{\mathrm{cell-out}} \quad (4.6)$$

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{(E_x^{\mathrm{miss}})^2 + \left(E_y^{\mathrm{miss}}\right)^2}. \quad (4.7)$$

The resolution of the $E_{x,y}^{\mathrm{miss}}$ components evaluated in simulation ranged from about 2 GeV (for a total transverse energy of 20 GeV) to 10 GeV (for a total transverse energy of 400 GeV) [122]. All changes to the objects (object energy scaling, smearing and systematic variations) that enter the $E_{\mathrm{T}}^{\mathrm{miss}}$definition were propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$object.

Chapter 5

Global sequential calibration

Contents

5.1 Overview

The determination of the jet energy scale (JES), its uncertainty and the achievement of an optimal jet resolution are major tasks in the ATLAS calibration program. ATLAS has developed several calibration schemes (see Section 4.7.2) to achieve these goals. In this section, the global sequential (GS) calibration scheme is presented in more details. GS is built upon a simpler jet calibration scheme (EM+JES) used for early analysis in ATLAS (see Section 4.7.4.1). GS is based on global jet observables that characterize the longitudinal and transverse jet structure and its main purpose is to improve the jet resolution without changing the average jet energy scale and to reduce the sensitivity of the calorimeter response to the jet flavour with respect to the EM+JES calibration.

This chapter is organized as follows. The details of the GS calibration technique and the derivation of the calibration constants using simulated events are discussed in Section 5.2. The data and event selection used for in this chapter is described in detail in Section 5.3. The performance of the GS calibration is discussed in Section 5.4. The validation of the calibration

using data is presented in Section 5.5. The evaluation of the systematic uncertainty associated to GS using inclusive multi-jets events and γ+jet events is described in Section 5.6.

5.2 Global sequential calibration description

The correction of the jet energy scale from the EM scale that brings the average jet response to 1 is done using the EM+JES calibration as described in Section 4.7.4.1. The EM+JES calibration provides a simple calibration that may be useful to evaluate systematic uncertainties for the early analyses but the jet resolution of this simple calibration for high-p_T jets is not as good as the one achieved by the GCW and LCW calibrations. In addition, the response to light-quark jets is 6% (3%) higher than the response to gluon jets at low (high) p_{T} in the barrel for the EM+JES calibration [123], which is a relatively big difference. The global sequential (GS) technique[1] uses few jet properties to improve the resolution and reduce the sensitivity of the response to jet flavor as compared to the EM+JES calibration.

The GS technique is a multi-variate extension of the EM+JES calibration. Any jet property x that carries information about the response of the jet can be used. A multiplicative correction to the jet energy is derived inverting the jet response as a function of this property:

$$C(x) = < \mathcal{R} >^{-1} (x). \tag{5.1}$$

After this correction, the remaining dependence of the response on the jet property x is removed without changing the average energy, resulting in a reduction of the spread of the reconstructed jet energy and, thus, an improvement in resolution. This is illustrated in Figure 5.1.

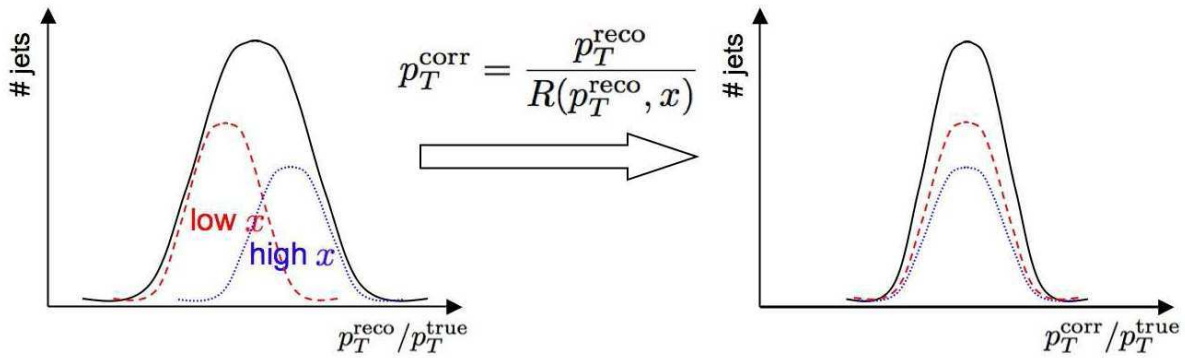


Figure 5.1: Illustration of the GS technique principle. The width of the jet response distribution (i.e. the jet resolution) is reduced when correcting the jet response dependence on x. Here p_{jet} refers to the four-vector representation of the jet [124].

The effect of the finite resolution of the calorimeter complicates the picture presented in Figure 5.1. In particular, since the correction is applied as a function of the transverse momentum of the calorimeter jet, $p_{\mathrm{T}}^{\mathrm{reco}}$, each jet in the dotted and dashed curves gets a different correction and the width of those curves changes in the correction process. For some properties, the dotted and dashed response distributions get wider during the correction process, then the improvement in the jet resolution is less pronounced [125]. GS uses only jet properties that provide the largest improvement in resolution, selected in an empirical way. Several jet properties can be used sequentially to achieve the optimal resolution. This procedure requires that the correction for

[1] Here, "global" refers to nature of the properties used, characterizing the jet as a whole and not its constituents.

jet property x_i (C^i) is calculated using jets to which the correction for jet property x_{i-1} (C^{i-1}) has already been applied. The four-vector representation of the jet after correction number i is given by :

$$p^i_{\text{jet}} = C^i(x_i) \times p^{i-1}_{\text{jet}} = C^i(x_i) \times C^{i-1}(x_{i-1}) \times p^{i-2}_{\text{jet}} = ...$$

This factorized procedure does not capture all the correlations between all properties. However, this effect has been shown to be very small and it doesn't degrade the improvement in resolution [124].

5.2.1 Monte Carlo determination of the GS corrections

In the GS context the jet response is calculated as:

$$\mathcal{R} = p^{\text{EM+JES}}_{\text{T,reco}} / p^{\text{truth}}_{\text{T}}, \tag{5.2}$$

where $p^{\text{EM+JES}}_{\text{T,reco}}$ corresponds to the transverse momentum of the calorimeter jet at the EM+JES scale and $p^{\text{truth}}_{\text{T}}$ to the transverse momentum of the corresponding matched truth jet[2]. The use of the jet p_{T} response instead of the jet energy response (as done for the EM+JES derivation) allows to validate directly the GS calibration in data using an in-situ technique. This technique exploits the transverse momentum balance in di-jet events leading to the estimation of an equivalent jet $p_{\text{T}}-$response (see Section 5.5 for more information).

First, the average jet response $\langle\mathcal{R}\rangle_{\text{truth}}$ is defined as the mean of a Gaussian fit to the peak of the response distribution $\mathcal{R}$ for each $(p^{\text{truth}}_{\text{T}}, \eta, x)-$bin. Then, $\langle\mathcal{R}\rangle_{\text{truth}}$ is used to extract the response as a function of $p^{\text{EM+JES}}_{\text{T,reco}}$ [115]:

$$\langle\mathcal{R}\rangle_{\text{reco}}(p^{\text{EM+JES}}_{\text{T,reco}}, \eta, x) = \langle\mathcal{R}\rangle_{\text{truth}}(\langle\mathcal{R}\rangle_{\text{truth}} \times p^{\text{truth}}_{\text{T}}, \eta, x). \tag{5.3}$$

Finally, the calibration constants are given by:

$$C(p^{\text{EM+JES}}_{\text{T,reco}}, \eta, x) = \langle R\rangle^{-1}_{\text{reco}}(p^{\text{EM+JES}}_{\text{T,reco}}, \eta, x). \tag{5.4}$$

Initially the corrections are derived in bins of $p^{\text{truth}}_{\text{T}}$, since the jet p_{T} response distribution is gaussian only in bins of $p^{\text{truth}}_{\text{T}}$. In bins of reconstructed p_{T}, the distribution is not gaussian due to the impact of the underlying p_{T} spectrum. By construction, this correction does not change the mean response of jets in the sample where it was calculated if the mean response was 1 before the correction. The effect of the correction on the mean response of jets from other samples is discussed in later sections.

5.2.2 Properties derived from the internal jet structure

The properties used in the GS calibration characterize the longitudinal and transverse structure of the jet. Indeed, such variables contain information about how the jet deposits energy in the calorimeter and thus what the calorimeter response to jets is. A large energy deposit in the hadronic layers indicates, for example, a larger hadronic component of the jet implying that, on average, the response of this jet will be low in the non-compensating ATLAS calorimeter. Close to a crack region, the transverse extension of the jet is correlated to how many particles of a jet hit the poorly instrumented transition region.

[2]Only calorimeter jets that are geometrically matched to truth jets within a cone of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.3$ around the jet axis are used for the calibration

Each of these jet properties may be sensitive to several effects: energy deposited in the dead material, non-compensation of the calorimeter or unmeasured energy due to the noise suppression. In the GS calibration, no attempt to separate all these effects is done. The jet properties that help to obtain a large improvement in resolution are used, implicitly correcting in average for those effects.

The longitudinal structure of the jet is characterized by the fractional energy deposited in the different longitudinal compartments of the calorimeters before any jet calibration is applied ("layer fractions" or f_{layer}) :

$$f_{\text{layer}} = \frac{E^{\text{EM}}_{\text{layer}}}{E^{\text{EM}}_{\text{jet}}}, \tag{5.5}$$

where $E^{\text{EM}}_{\text{jet}}$ is the jet energy at the EM scale and $E^{\text{EM}}_{\text{layer}}$ the energy deposited in the layer of interest, also defined at the EM scale. The fractional energy is used instead of the energy in the layer to reduce the $p^{\text{jet}}_{\text{T}}$ dependence of the distribution, but comparable results have been obtained in the Monte Carlo simulation using the energies in the different layers [125]. The transverse structure can be characterized by the jet width, defined as the average distance between the jet constituents and the jet axis, weighted by the transverse constituent momentum p_{T}:

$$\text{width} = \frac{\sum_{\text{i}} p^{i}_{\text{T}} \Delta R_{i,\text{jet}}}{\sum_{\text{i}} p^{i}_{\text{T}}}, \tag{5.6}$$

where the sums are over the jet constituents (i) and $\Delta R_{i,\text{jet}} = \sqrt{(\eta^{\text{jet}} - \eta^{\text{cluster } i})^2 + (\phi^{\text{jet}} - \phi^{\text{cluster } i})^2}$ is the distance in $(\eta \times \phi)-$space between the jet constituents and the jet axis. The different types of jet constituents were described in Section 4.7.1.1. In the studies presented in this chapter, topo-clusters are used as jet constituents.

The properties used as well as the order in which they are applied are summarized in Table 5.1. The fraction of the EM scale jet energy deposited in the first layer of the Tile calorimeter is denoted as $f_{\texttt{Tile0}}$; the one deposited in the first layer of the forward calorimeter $f_{\texttt{FCal1}}$; the one deposited in the presampler f_{PS} and the one deposited in the third layer of the EM calorimeter $f_{\texttt{LAr3}}$. Finally, width refers to the jet width. The improvement in resolution obtained is independent of which property is used first to derive a correction [125].

$\lvert\eta\rvert$ region	Corr 1	Corr 2	Corr 3	Corr 4
$\lvert\eta\rvert < 1.2$	$f_{\texttt{Tile0}}$	$f_{\texttt{LAr3}}$	f_{PS}	width
$1.2 \leq \lvert\eta\rvert < 1.4$	$f_{\texttt{Tile0}}$			width
$1.4 \leq \lvert\eta\rvert < 1.7$	$f_{\texttt{Tile0}}$	$f_{\texttt{HEC0}}$		width
$1.7 \leq \lvert\eta\rvert < 3.0$		$f_{\texttt{HEC0}}$		width
$3.0 \leq \lvert\eta\rvert < 3.2$		$f_{\texttt{LAr3}}$		width
$3.2 \leq \lvert\eta\rvert < 3.4$		$f_{\texttt{LAr3}}$		
$3.4 \leq \lvert\eta\rvert < 3.5$		$f_{\texttt{LAr3}}$		width
$3.5 \leq \lvert\eta\rvert < 3.8$	$f_{\texttt{FCal1}}$			width
$3.8 \leq \lvert\eta\rvert < 4.5$	$f_{\texttt{FCal1}}$			

Table 5.1: Sequence of corrections in the GS calibration scheme in each $|\eta|$ region.

In the rest of this thesis, "GSL" will refer to the calibration applied up to the third correction (that is containing only the layer fractions corrections) and "GS" to the calibration applied up

to the last correction (including the width correction). The full jet calibration procedure for jets calibrated with the GS calibration is pictured in Figure 5.2.

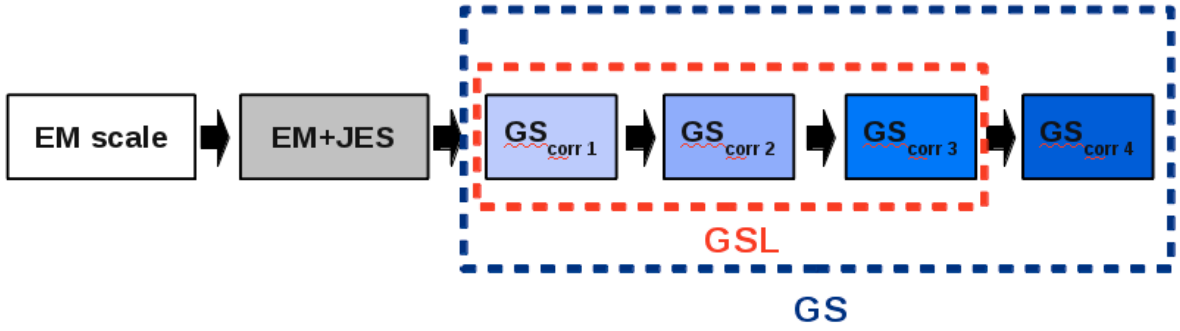


Figure 5.2: Schematic representation of the jet calibration procedure for jets calibrated with the GS calibration schemes.

5.2.3 Technical details

As an illustration of how the derivation of the GS corrections works in detail, we present in this section the derivation for jet property $f_{\texttt{Tile0}}$. The derivation of the correction for the other jet properties is similar.

- The GS corrections are determined in 45 jet $|\eta|$ bins of width 0.1 from $|\eta| = 0$ to $|\eta| = 4.5$. The original detector η is used instead of the origin-corrected η_{origin}, since the first one corresponds directly to a region of the calorimeter.

- In each η bin, 20×25 $(x_i, p_{\text{T}}^{\text{truth}})-$bins are defined. The $p_{\text{T}}^{\text{truth}}$ bins are limited by the following values in GeV: 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 350, 400, 450, 500, 600, 700, 800, 1000, 1200. The x_i bins are of variable size (determined such that each bin has the same number of jets).

- The jet response distribution in each $(x_i, p_{\text{T}}^{\text{truth}})-$bin is fitted with a gaussian as illustrated in Figure 5.3. The center of the $p_{\text{T}}^{\text{truth}}$ bins, the center of the x_i bins and the mean of the gaussian fits , $\langle\mathcal{R}\rangle(p_{\text{T}}^{\text{truth}}, x_i)$, are used to built a two-dimensional graph of the jet response as a function of x_i and $p_{\text{T,reco}}^{EM+JES}$ (for each of the points in the graph, the value of the $p_{\text{T}}^{\text{truth}}$ coordinate is changed to $p_{\text{T,reco}}^{EM+JES} = p_{\text{T}}^{\text{truth}} \times \langle\mathcal{R}\rangle(p_{\text{T}}^{\text{truth}}, x_i)$). There is one two-dimensional graph for each bin of $|\eta|$.

- These jet response graphs can be interpolated or fitted to calculate the calibration constants in Equation 5.4. The interpolation is simpler to implement, since no knowledge of the fit function is needed. Moreover, the number of events in each $(|\eta|, p_{\text{T}}^{\text{truth}}, x_i)-$ bin is sufficient so that statistical fluctuations are not a limiting factor. The interpolation technique was therefore used in all the studies presented in this document. An example of an interpolated graph is shown on Figure 5.4.

- Only calorimeter jets that are geometrically matched to truth jets within a cone of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.3$ around the jet axis are used for the calibration.

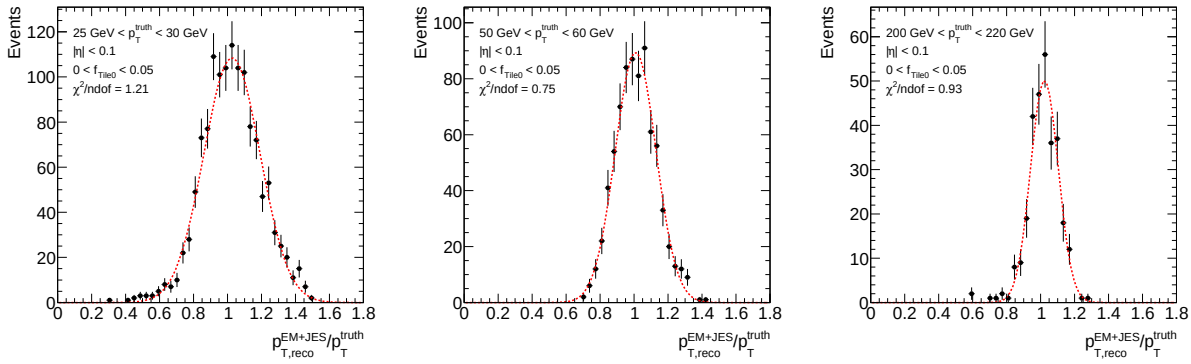


Figure 5.3: Gaussian fits of the jet response distribution in different $p_{\mathrm{T}}^{\mathrm{truth}}$ bins: (a) $25 < p_{\mathrm{T}}^{\mathrm{truth}} < 30$ GeV, (b) $50 < p_{\mathrm{T}}^{\mathrm{truth}} < 60$ GeV and (c) $200 < p_{\mathrm{T}}^{\mathrm{truth}} < 220$ GeV, for jets reconstructed from topoclusters with the anti-k_t algorithm with $R = 0.6$.

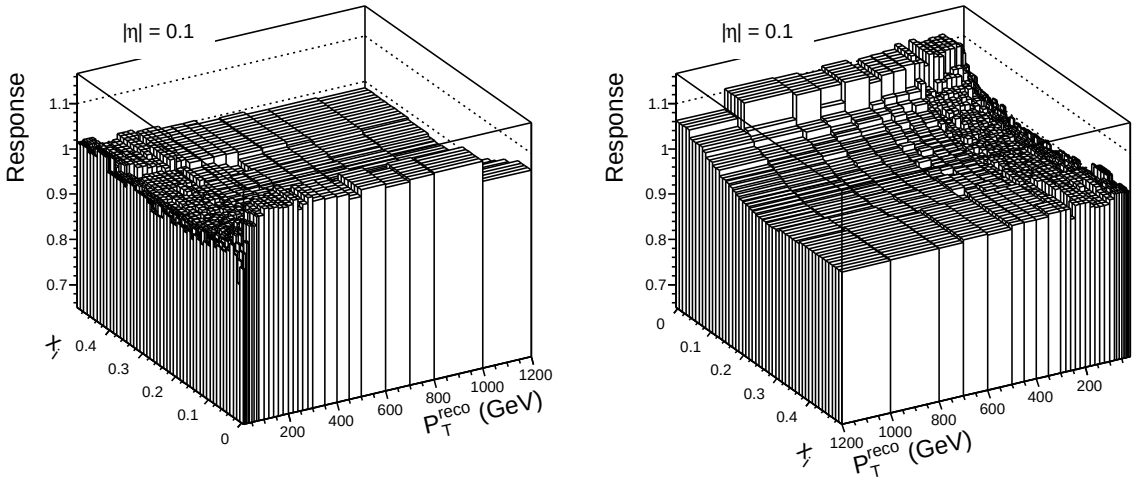


Figure 5.4: Interpolation of the two-dimensional graph of the jet response as a function of x_i and $p_{\mathrm{T,reco}}^{\mathrm{EM+JES}}$ for jets with $|\eta| < 0.1$ reconstructed from topoclusters with the anti-k_t algorithm with $R = 0.6$.

- Due to file size constraints the jet reconstruction was configured to only keep jets with $p_{\mathrm{T,reco}}^{\mathrm{EM}} > 7$ GeV. It limits the information of very low $p_{\mathrm{T}}^{\mathrm{truth}}$ jets as some of the reconstructed jets might fall below this threshold. In order to avoid any possible bias in the jet response due to the reconstruction threshold the gaussian fits were only performed for $p_{\mathrm{T}}^{\mathrm{truth}} > 15$ GeV. All jets below 15 GeV were calibrated with the calibration constant derived for 15 GeV jets.

- In order to avoid cases were one particle jet is splitted into two reconstructed jets or two particle jets are merged into one reconstructed jet, only isolated jets are used to derived the calibration constants. An isolated reconstructed (truth) jet is defined as a jet that has no other reconstructed (truth) jet with EM+JES scale (truth) $p_{\mathrm{T}} > 7$ GeV within

$\Delta R = 2.5R$, where R is the distance parameter of the jet algorithm.

- The GS calibration constants were derived for jets reconstructed using the anti-k_t algorithm for two different distance parameters $R = 0.4$ and 0.6 and two different calorimeter inputs (towers and topoclusters). The derivation of the corrections for a given reconstruction algorithm, distance parameter and type of calorimeter input takes around one day (it includes the time to cross check the results of the different fits and interpolation plots). They were included in the official calibration tool used in the ATLAS collaboration JetCalibTools[3] and are available in the D3PDs produced by the ATLAS JetEtMiss group. In addition, a standalone C++/ROOT class has been created to apply GS calibration constants in analyses outside the ATHENA framework[4].

5.3 Data and event selection for 2010 data

5.3.1 Data sample

Proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV, recorded from March to October 2010 are used. The data set corresponds to an integrated luminosity of 35 pb^{-1}. The evolution of the integrated luminosity as a function of time for the period when these data were collected is shown in Figure 2.2. Data from 2010 is divided into periods going from A to I. Events collected using the minimum bias and calorimeter triggers were required to belong to specific luminosity blocks in which the calorimeters, the inner detector, the solenoid, the trigger and the reconstructed physics objects have passed a data-quality assessment and are deemed suitable for physics analysis (see Section 4.1 for more information about the data quality requirements).

5.3.2 Monte Carlo simulation

Data are compared to several Monte Carlo simulations of non-diffractive pp collisions at 7 TeV. An overview of Monte Carlo generators for LHC physics can be found in Ref. [69].

Inclusive QCD jet events were generated with Pythia [78], and Herwig++ [82]. Pythia implements leading-order matrix elements from perturbative QCD for 2→2 processes, followed by parton showers to model additional radiation in the leading-logarithmic approximation. Hadronization, fragmentation and soft multiple-parton interactions are also simulated within Pythia. The Pythia generator has been validated with data collected by the ATLAS detector at $\sqrt{s} = 0.9$ and 7 TeV [93, 1]. Two sets of event samples were generated using Pythia:

- ATLAS MC10 tune [126] (also known as AMBT1): it uses the charged particle spectra measured by ATLAS in minimum bias collisions for tuning multiple parton interactions. About 12.6 millions of di-jet events were generated using this tune. It corresponds to the nominal (baseline) Monte Carlo sample in this study.

- Perugia2010 tune [127]: it has increased final state radiation to better reproduce the jet shapes and hadronic event shapes using LEP and Tevatron data. About 3.6 millions of events were generated with this tune.

Herwig++ has similar leading order matrix elements as Pythia, but uses an angular-ordered parton shower and a cluster hadronization model. For the Herwig++ sample about 2.8 millions of events were generated. For the study of pile-up, two samples generated also with Pythia have been used, one for in-time and one for out-of-time pile-up. The first one simulates

[3] http://alxr.usatlas.bnl.gov/lxr/source/atlas/Reconstruction/Jet/JetCalibTools/JetCalibTools
[4] https://twiki.cern.ch/twiki/bin/view/Main/ResultsGSC#Using_the_Global_Sequential_Cali

additional proton-proton interactions per bunch crossing while the second one contains pile-up arising from the proximity of bunches and includes effects from the difference in bunch spacing between bunches within bunch train and the first and the last bunch in the train (bunch-train pile-up). The in-time pile-up samples has about 2.5 millions of events, while the out-of-time pile-up sample has about 12.1 millions of events.

The generator events are passed through the standard GEANT4 [91, 92] simulation of the ATLAS detector and then fully reconstructed and analyzed with the same software as used for the data processing. The ATLAS detector geometry used in the simulation of the nominal sample reflects the geometry of the detector as best known at the time of these studies.

5.3.3 Event selection

The same event selection criteria are used in the Monte Carlo simulation and in the data, except the trigger selection for data-driven studies presented in Section 5.5, which was only applied in data. Events are required to have at least one vertex with 5 or more tracks with $p_{\mathrm{T}}^{\mathrm{track}} > 150$ MeV. The longitudinal position of the vertex is required to be within 10 cm of the detector center. These cuts serves to reject events originating from cosmic rays and particles produced in interactions of the beam with particles in the beam tunnel ("beam halo" and "beam gas" events).

The trigger selection for data-driven studies presented in Section 5.5 uses either the level 1 minimum bias trigger or the level 1 and level 2 calorimeter triggers depending on the data period and the p_{T} range of interest. To cover the region $p_{\mathrm{T}} < 40$ GeV, events from the minimum bias stream were used if there was at least one hit in the Minimum Bias Trigger Scintillators (MBTS). Events from the calorimeter trigger stream were required to satisfy a central single Level-1 jet trigger for early data periods and a central single High Level jet trigger for later data periods, since the trigger selection and prescalings change during data taking. The Level 1 central single jet trigger accepts the event if the EM scale energy of any jet (defined in term of RoIs as described in Section 2.2.6) in the event with $|\eta| \leq 3.2$ is above a certain threshold. High Level central single triggers are seeded by a Level 1 central single trigger and also accept events where at least one jet has an EM scale energy above a given threshold. The event trigger selection follows that described in Table 5.2. It was chosen such that the trigger efficiency, for a specific region of p_{T}, was greater than 99% and approximately flat as a function of the jet η. The functioning and performance of the jet triggers in the first 2010 data is described in Ref. [128].

p_{T} region (GeV)	Period A → F	Period G → I
20-30	MinBias	MinBias
30-40	MinBias	MinBias
40-50	L1_J5	EF_j20_jetNoEF
50-60	L1_J10	EF_j30_jetNoEF
60-80	L1_J15	EF_j35_jetNoEF
80-110	L1_J15	EF_j35_jetNoEF
110-160	L1_J30	EF_j50_jetNoEF
160-210	L1_J55	EF_j70_jetNoEF
210-260	L1_J75	EF_j95_jetNoEF
260-310	L1_J95	L1_J95
310-400	L1_J95	L1_J95
400-800	L1_J95	L1_J95

Table 5.2: Triggers used for each p_{T} region in the central region of the ATLAS detector $|\eta| < 2.8$.

5.3.4 Jet reconstruction and selection

Calorimeter jets are reconstructed using the anti-k_t jet algorithm [106] with a distance parameter $R = 0.6$ using the FastJet software [129]. Eventhought the GS corrections have been derived for both types of jet constituents (topoclusters and towers), the studies presented in this chapter were done using topological clusters at the EM scale as inputs to the jet reconstruction algorithm. In the Monte Carlo simulation, truth jets are reconstructed using the same algorithm as calorimeter jets. Only jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV and $|\eta| < 2.8$ are used in studies of performances (for the derivation of the GS calibration jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 15$ GeV and $|\eta| < 4.5$ were used).

Jets must also pass several selection criteria which are each designed to mitigate the impact of specific non-collision backgrounds. These criteria are only applied to data and are discussed in detail in Ref. [130]. These *jet cleaning cuts* have an efficiency above 99%. They target three main category of non-collision backgrounds:

Removal of coherent noise in the EM calorimeter

- The cell signal quality factor (f_{QLar}), representing the fraction of cells with a poor signal quality defined by the pulse shape must be smaller than 0.8 for jets which deposit at least 95% of their energy in the electromagnetic calorimeter.

Removal of hadronic end-cap calorimeter sporadic noise

- The absolute value of the apparent negative energy of the neighboring calorimeter cells (E_{neg}) produced due to the capacitive coupling between channels must be smaller than 60 GeV.

- The cell signal quality factor (f_{QLar}) must be smaller than or equal to 0.5 for jets which deposit more than 50% of their energy in the HEC ($f_{\mathrm{HEC}} > 0.5$).

Removal of cosmic and beam Background

- The absolute value of the energy-squared-weighted cell time of the jet, known as jet timing, must be less than or equal to 10 ns.

- At least 5% of the jet energy must be deposited in the electromagnetic calorimeter for jets with $|\eta| < 2.0$.

- The fraction of the jet energy deposited in one of the calorimeter layers cannot be bigger than 99% for jets that fall on the central region ($|\eta| < 2.0$).

- The ratio of the scalar sum of the p_{T} of the tracks pointing towards the jet and the calibrated jet p_{T} (f_{ch}) must be at least 5%, for central jets ($|\eta| < 2.0$) which deposit more than 95% of their energy in the electromagnetic calorimeter.

5.4 Performances of the GS calibration

The performance of the GS calibration has been extensively studied in Monte Carlo simulations [125, 124]. In this section the most relevant results regarding the response linearity, the estimation of the jet resolution improvement, the sensitivity to jet flavour and to the presence of pile-up are presented.

5.4.1 Jet response linearity and jet transverse momentum resolution

The jet p_{T} response is defined in $(p_{\mathrm{T}}^{\mathrm{truth}}, |\eta|)-$ bins as:

$$\mathcal{R} = p_{\mathrm{T,reco}}^{\mathrm{calibrated}} / p_{\mathrm{T}}^{\mathrm{truth}}, \tag{5.7}$$

where the reconstructed jet is calibrated using a given calibration scheme. For example the p_{T} response for the GSL calibration will be calculated using reconstructed jets calibrated using the GSL calibration, i.e $p_{\mathrm{T,reco}}^{\mathrm{GSL}}$. The jet transverse momentum resolution (σ_{R}/R) is defined as the ratio of the width of the jet p_{T} response by its average. Figure 5.5 shows the evolution of the jet response and resolution as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ in the nominal Pythia Monte Carlo simulation (see Section 5.3.2) where the calibration was derived as the different GS corrections are applied in different regions of the detector. The GCW and LCW calibrations are also shown for reference. The response after each correction is applied remains almost unchanged as expected, except for jets in the first bin with $p_{\mathrm{T}}^{\mathrm{truth}} < 30$ GeV where a difference of 1-1.5% is observed between the jet response at the EM+JES and GS scale.

The response linearity can also be verified as a function of the different jet properties used to derive the GS corrections (see Table 5.1). Figure 5.6 shows the mean of the jet response distribution in Monte Carlo simulation after the EM+JES calibration has been applied as a function of some representatives jet properties used in the GS calibration [2]. For all jet properties used to calculate the GS corrections a strong dependence of the response as a function of the property is observed. Four representatives $p_{\mathrm{T}}^{\mathrm{truth}}$ bins are shown. The same distributions after the GS calibration are shown in Figure 5.7. The jet response is roughly 1 for all the jet properties used in the correction after applying the GS calibration.

The jet resolution improves as we add more corrections. After the four corrections, the resolution obtained is comparable to that of the jets calibrated using the other calibration schemes. A more quantitative estimate can be obtained evaluating the relative improvement in the jet transverse momentum resolution with respect to the EM+JES calibration:

$$\delta(\sigma_{\mathrm{R}}/R) = \frac{(\sigma_{\mathrm{R}}/R)^{\mathrm{EM+JES}} - (\sigma_{\mathrm{R}}/R)}{(\sigma_{\mathrm{R}}/R)^{\mathrm{EM+JES}}} * 100. \tag{5.8}$$

To do it σ_{R}/R is calculated as a function of p_{T} through a fit of the form:

$$\frac{\sigma_{\mathrm{R}}}{R} = \frac{S}{\sqrt{p_{\mathrm{T}}}} \oplus \frac{N}{p_{\mathrm{T}}} \oplus C, \tag{5.9}$$

where $\oplus$ is used to indicate addition in quadrature and N, S and C are the noise, stochastic and constant terms, respectively. N parametrizes fluctuations due to electronic noise and offset energy from multiple interactions, S parametrizes the stochastic fluctuations in the amount of energy sampled from the jet hadron shower and C encompasses the fluctuations that are a constant fraction of the energy as contributions due to detector imperfections and leakage. The relative improvement in the transverse momentum resolution with respect to EM+JES for four representative jet p_{T} and for $|\eta| < 0.3$ is summarized in Table 5.3. Table 5.3 shows that the jet transverse momentum resolution obtained after GS is comparable to the one obtained by using GCW and LCW for jets with $|\eta| < 0.3$. The improvement in the resolution is similar $0.3 < |\eta| < 1.2$. For $1.2 < |\eta| < 4.5$ the improvement with respect to the resolution for EM+JES is slightly smaller that the one observed in the central region of the detector, but still comparable with the improvement achieved when using LCW or GCW as can be seen in Appendix A.

It is also worthwhile mentioning that Figure 5.5 shows that in general the resolution is better when going to higher η regions, since η determines the energy of a jet of a given p_{T}:

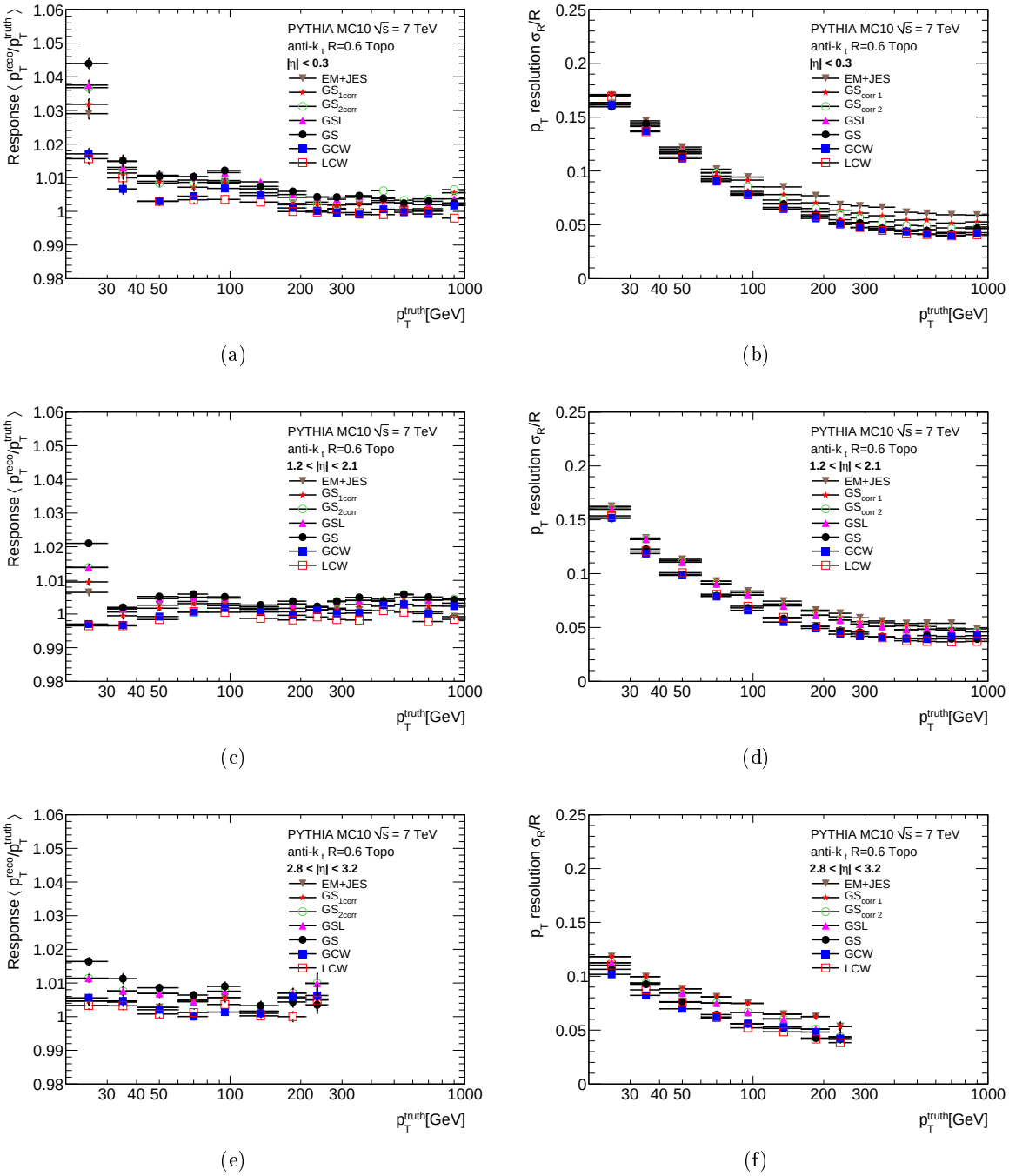


Figure 5.5: Jet response (left) and jet transverse momentum resolution (right) as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ for different $|\eta|$ regions after each GS correction in the nominal MC simulation for jets reconstructed with the anti-k_t algorithm $R = 0.6$. GCW and LCW results are also shown for comparison.

$$E \approx p_{\mathrm{T}}/\cosh|\eta|. \tag{5.10}$$

Therefore, as the jet energy increases while going to the more forward bins, the resolution improves. The jet resolution was also measured for jets reconstructed using the anti-k_t algorithm

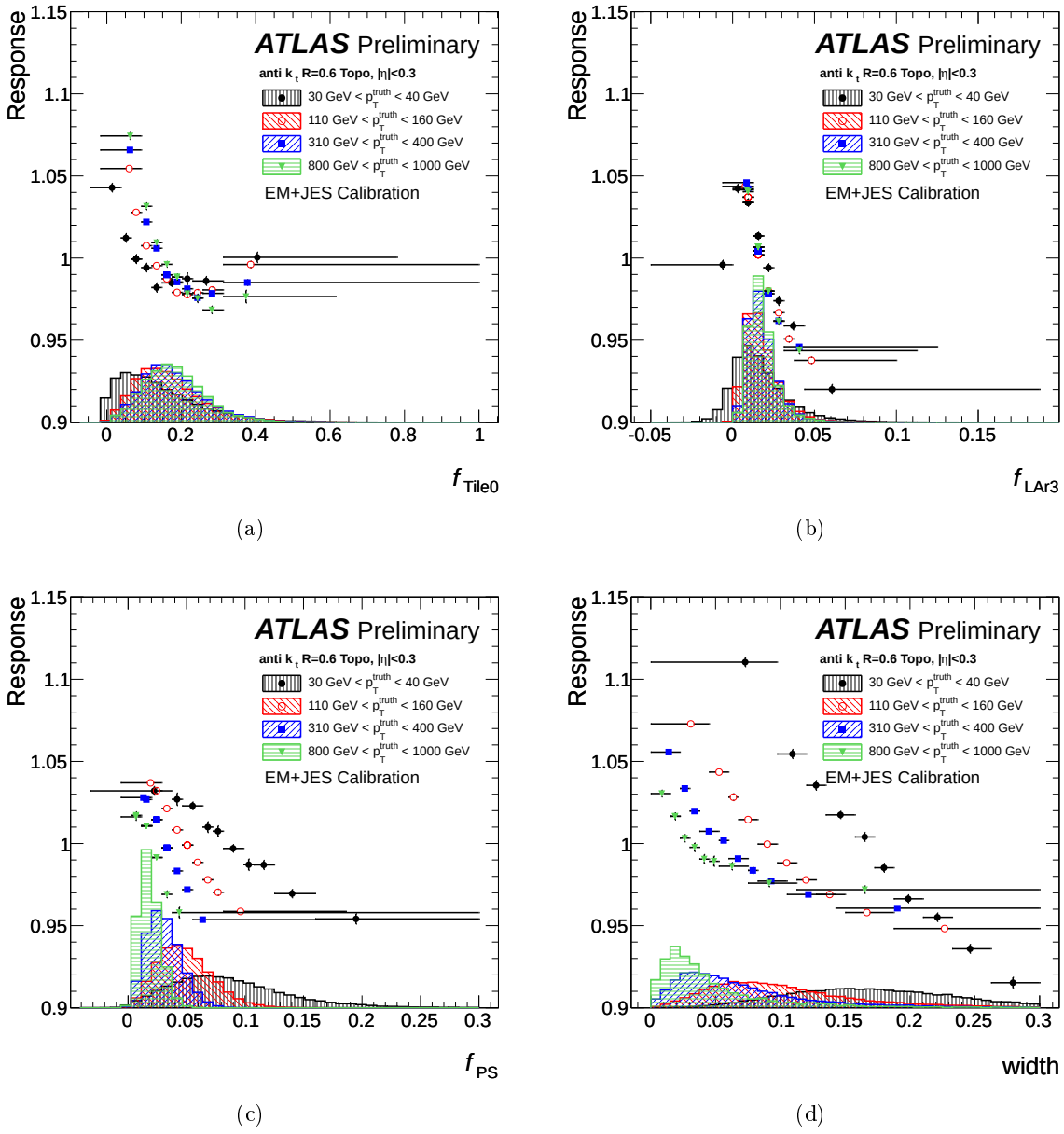


Figure 5.6: Jet energy response as a function of (a) $f_{\texttt{Tile0}}$, (b) $f_{\texttt{LAr3}}$, (c) f_{PS} and (d) the jet width obtained with Monte Carlo simulation using jets with $|\eta| < 0.3$ for different bins of matched particle jet $p_{\mathrm{T}}^{\mathrm{truth}}$. Jets are reconstructed with the anti-k_t algorithm $R = 0.6$. All jets have been calibrated using the EM+JES calibration. The underlying distributions of the corresponding jet properties are also shown.

with a distance parameter $R = 0.4$. It was found to be slightly smaller than the one in Figure 5.5. This can be attributed to the fact that out of cone losses are more important for $R = 0.4$ jets.

The resolution improvement has been also measured in data using data-driven techniques [112]. The jet transverse momentum resolution as a function of the average jet transverse momenta for events with two jets in the same rapidity bin ($|y| < 0.8$) is shown in Figure 5.8 for the four jet calibration schemes (those have been also fitted using Eq. 5.9). The relative improvement in the jet transverse momentum resolution as a function of the average jet trans-

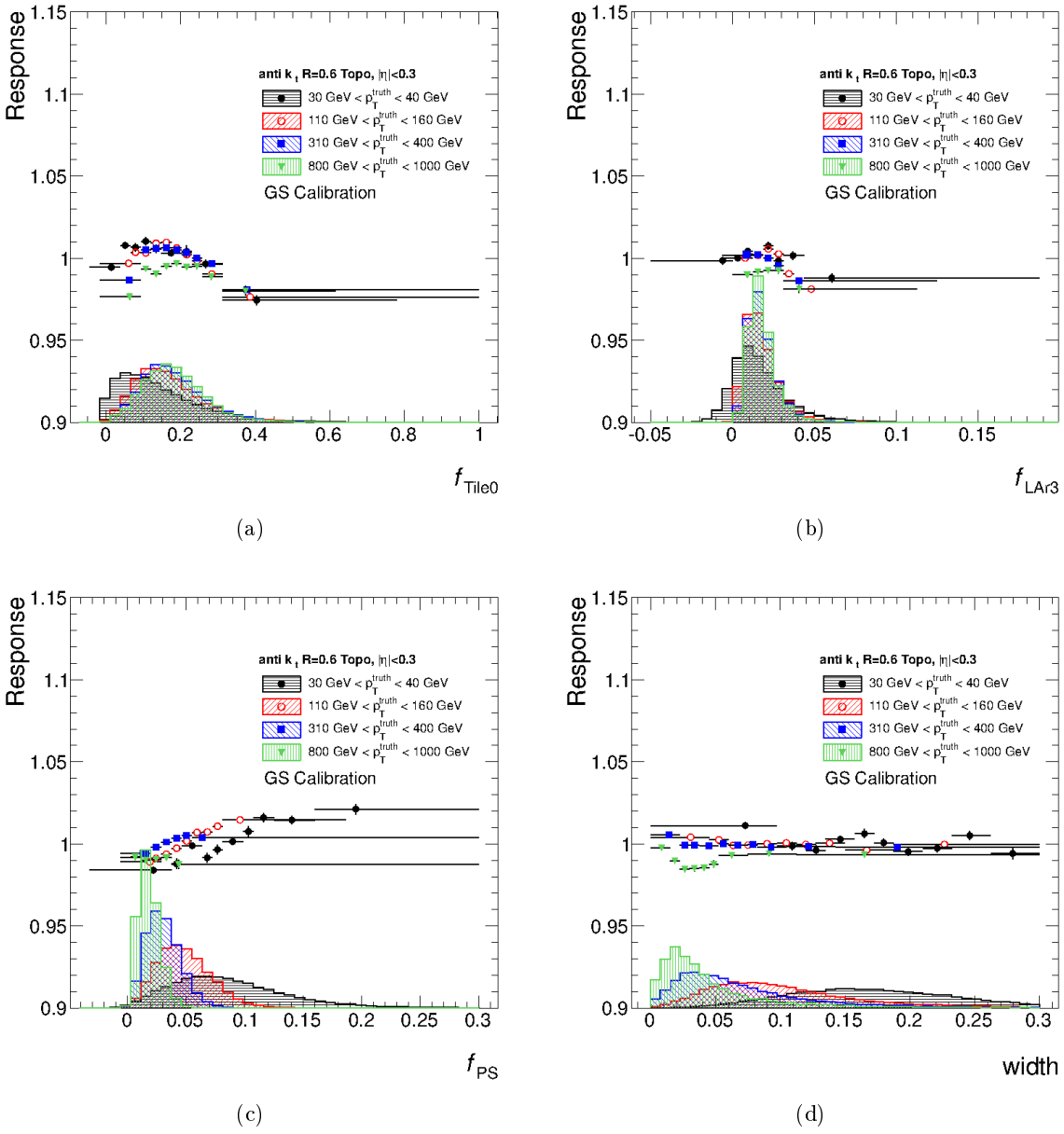


Figure 5.7: Jet energy response as a function of (a) $f_{\texttt{Tile0}}$, (b) $f_{\texttt{LAr3}}$, (c) f_{PS} and (d) the jet width obtained with Monte Carlo simulation using jets with $|\eta| < 0.3$ for different bins of matched particle jet $p_{\mathrm{T}}^{\mathrm{truth}}$. Jets are reconstructed with the anti-k_t algorithm $R = 0.6$. All jets have been calibrated using the GS calibration. The underlying distributions of the corresponding jet properties are also shown.

verse momentum taking the EM+JES as baseline is shown in Figure 5.8 and is summarized for three representative average jet p_{T} in Table 5.4. After GS calibration the relative improvement in resolution using data is comparable to that of GCW and LCW. The improvement is found to be up to 30 % at 400 GeV for GSC, LCW and GCW. Results presented in Tables 5.3 and 5.4 are not directly comparable: one is done in η bins while the other in done in y bins, the binning is different, no isolation cut is applied in resolution calculated using data-driven techniques and all isolated jets in the events are used for Figure 5.5, while only the two leading jets are considered

Calibration	$\delta(\sigma_R/R)$ at 35 GeV (%)	at 150 GeV (%)	at 500 GeV (%)	at 1000 GeV (%)
GS_{corr1}	1.1	8.3	9.8	10.6
GS_{corr2}	3.6	14.2	19.2	18.6
GSL	3.8	17.8	24.7	20.9
GS	3.4	18.6	26.4	20.8
GCW	6.4	24.1	31.5	27.6
LCW	6.9	22.4	32.3	30.8

Table 5.3: Relative jet resolution improvement after GS, GCW and LCW in the nominal Monte Carlo simulation for jets with $|\eta| < 0.3$.

for Figure 5.8. However, both studies shows an improvement in the jet resolution resolution that increases with increasing p_T and which is comparable with the improvement achieved when using GCW and LCW.

Calibration	$\delta(\sigma_R/R)$ at 35 GeV (%)	at 150 GeV (%)	at 500 GeV (%)
GS	14	26	32
GCW	15	28	30
LCW	12	27	30

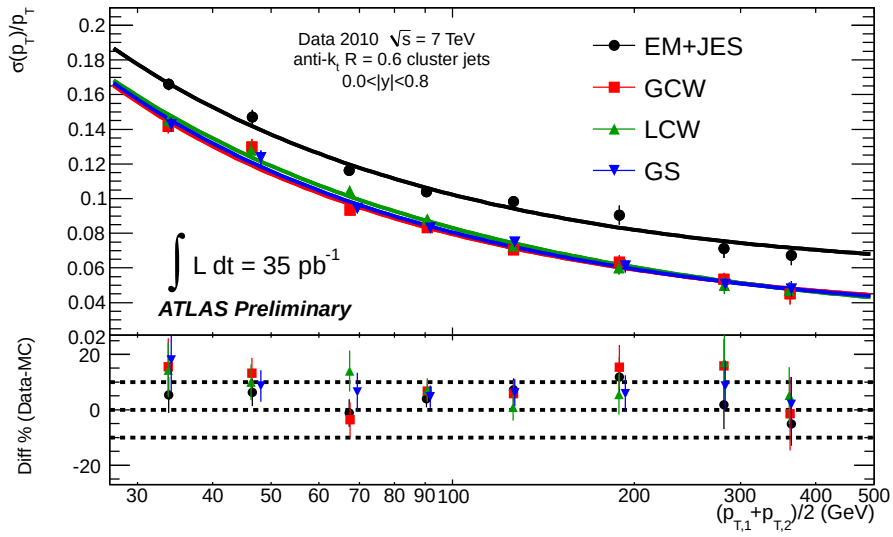
Table 5.4: Relative jet resolution improvement after GS, GCW and LCW using data for jets with $|y| < 0.8$.

5.4.2 Flavour dependence of the jet response

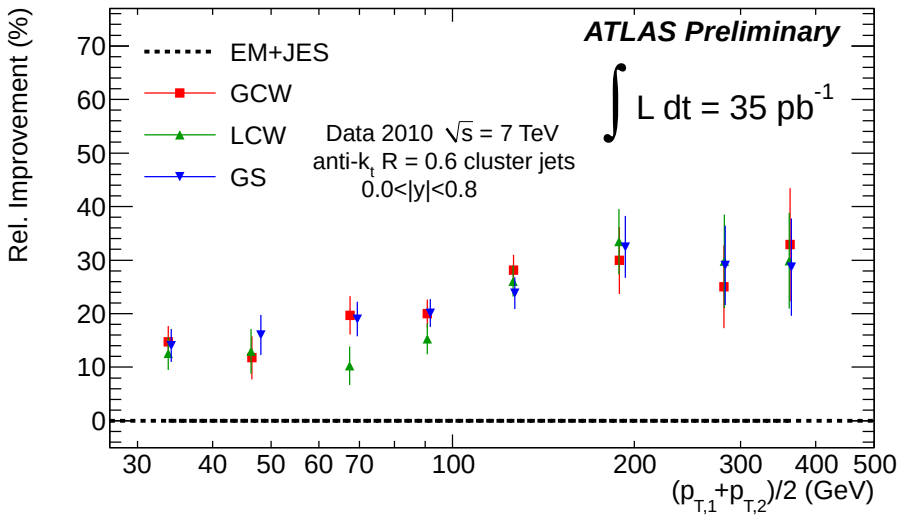
It has been shown in Section 5.4.1 that the GS calibration preserves linearity in the nominal Monte Carlo sample used to derive the corrections. If the calibration is applied to a sample with jets coming from a different physics process the mean response will be different from 1. The jet properties and thus the calorimeter response are sensitive to differences in fragmentation between quark and gluon initiated jets produced in different physics processes.

The difference in jet response between gluon-initiated and light-quark-initiated (LQ) jets have been studied in detail in [123, 3] for the different jet calibration schemes. Gluon- and LQ-jets are obtained from Monte Carlo multijet events throught a geometric matching with the partons. The flavour dependence of the jet response is in part a result of the differences in particle level properties of the two types of jets. A gluon-jet tends to have more particles, and those particles tend to be softer than in the case of a LQ-jet. Also, a gluon-jet tends to be wider (i.e. with lower energy density in the core of the jet) before interacting with the detector. The magnetic field in the inner detector amplifies the broadness of the gluon-jet, since their low-p_T charged particles tend to bend more than the higher p_T particles in a LQ-jet. The harder particles in a LQ-jet tend to penetrate further into the calorimeter.

Figure 5.9 shows the flavor dependence of the average jet response for jets calibrated with the EM+JES, GS, GCW and LCW calibrations. For jets calibrated with the EM+JES scheme the difference between the LQ-jet and gluon-jet average response ranges between 5-6% at low p_T. This difference decreases to around 2-3% at high p_T. The difference between LQ-jet and gluon-jet average response is correlated with differences in the jet properties. Therefore, more complex jet calibration schemes that account for jet shower properties variations can partially correct the flavour dependence. For jets calibrated with the LCW and GCW calibrations LQ-jets



(a)

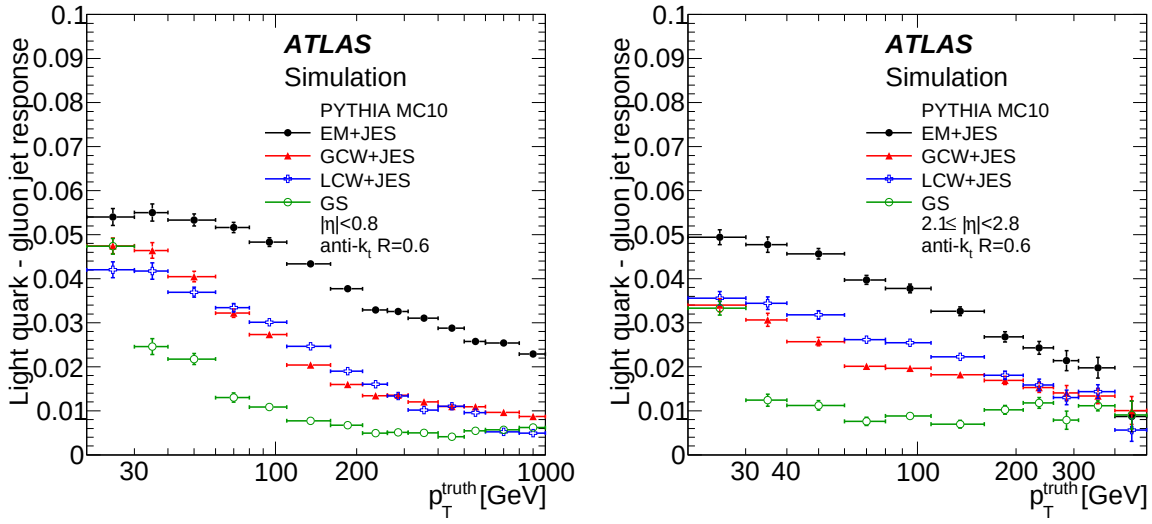


(b)

Figure 5.8: Jet transverse momentum resolution (a) and relative improvement in the jet transverse momentum resolution (b) as a function of the average jet transverse momenta for real data events with two jets in the same rapidity bin ($|y| < 0.8$) for EM+JES, GCW, LCW and GS calibrations. In Figure (a) the lines correspond to the fits for each JES scheme. The lower plot shows the relative difference between Monte Carlo and the data results (using the same data driven technique). In Figure (b) the EM+JES jet calibration scheme is taken as baseline (black dotted line) [112].

have 4-5% higher average jet response than gluon jets at low p_{T} and 1-1.5% at high p_{T}.

In the case of jets at the GS scale the flavour dependence of the average jet response is reduced to about 2-3% at low p_{T} and below 1% at high p_{T}. In particular, the width correction in GS, has proved to be helpful to reduce the flavor dependence of the jet energy scale not only when it is applied in the GS context but also when it is derived and applied after GCW and LCW, as the transverse structure of the jet is very sensitive to the jet flavor (see [131]).



(a) (b)

Figure 5.9: Difference in average response of gluon- and LQ-jets as a function of the particle jet $p_{\mathrm{T}}^{\mathrm{truth}}$ for anti-k_t jets with $R = 0.6$ in (a) the barrel and (b) the end-cap in Monte Carlo simulation for different jet calibration schemes: EM+JES, GS, GCW and LCW [123].

5.4.3 Sensitivity to pile-up

The LHC in nominal conditions will generate up to an average of 23 minimum bias interactions together with the hard-scatter process of interest. Then it is important to check the robustness of the GS calibration when applied in the presence of pile-up. This robustness translates into small variations in the size of each of the corrections and the distributions of the jet properties. This makes the corrections derived in the sample without pile-up applicable to the sample with pile-up with a small effect on the jet energy scale.

Figure 5.10 shows the difference between the average response after each GS correction and the average response after the EM+JES calibration in the Monte Carlo simulation samples with in-time and bunch-train pile-up described in Section 5.3.2. The additional number of primary vertices in the in-time (bunch-train) pile-up sample is 1.7 (1.9) on average. The average response after each of the corrections is degraded by less than 1% for all the $p_{\mathrm{T}}^{\mathrm{truth}}$ bins except for $20 < p_{\mathrm{T}}^{\mathrm{truth}} < 30$ GeV where degradations of 2.5% are observed. These variations are smaller than the uncertainty on the jet energy in the absence of pile-up over the entire p_{T} range, thus demonstrating the robustness of GS with respect to pile-up. This conclusion holds for the pile-up conditions in 2010. In 2011 the number of pile-up interactions increases, events with more that 20 vertices were observed. This study should be repeated in order to determine the robustness of GS for high luminosity conditions.

5.5 Validation of the GS calibration using data

All studies performed show that the GS calibration preserves the energy scale of the EM+JES calibration scheme in the inclusive reference Pythia sample used to derive the corrections. However, a good performance of the calibration in the inclusive reference Pythia sample does not guarantee a similarly good performance in other data or Monte Carlo samples with different

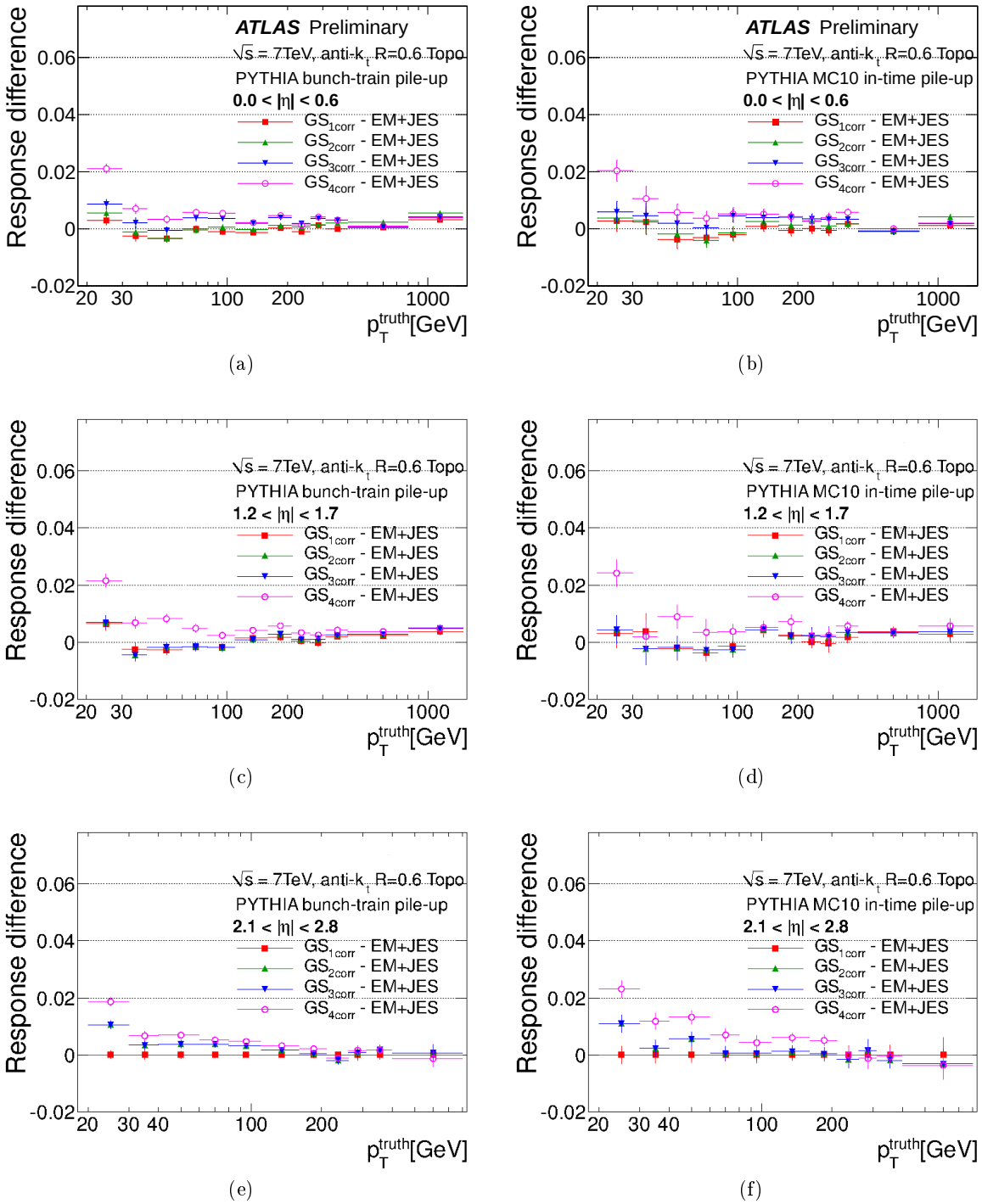


Figure 5.10: Difference between the average response after each GS correction and the average response after the EM+JES calibration as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ in samples with bunch-train pile-up (left) and in-time pile-up (right) for different $|\eta|$ regions [2].

topologies and flavor composition. The data-driven method that will be presented in this section allows to calculate the GS corrections from data. The comparison between the data-based and the Monte Carlo-based GS corrections allows to validate the last ones.

5.5.1 Di-jet balance method

The GS corrections can be derived from di-jet events using the di-jet balance method. This method is a tag and probe technique exploiting the p_{T} imbalance between two back-to-back jets. The version of the method used differs from those used in other studies in ATLAS [112, 132] in that it does not assume that the two jets balance at the truth jet level, but includes instead a correction for a potential truth jet imbalance, which will be described below.

Di-jet events are selected by requiring that the two highest p_{T} jets are back-to-back. The two jets are required to be in the same pseudorapidity region. Although this requirement reduce the statistics for $|\eta| > 2.8$, it allows comparing directly the p_{T} imbalance in data and Monte Carlo in a given pseudorapidity region avoiding contributions from the η inter-calibration differences, which were found to go up to 3% for $p_{\mathrm{T}} > 60$ GeV and to 12% for $p_{\mathrm{T}} < 60$ GeV, in the region $|\eta| < 2.8$ [132]. The jet whose response dependence on the layer fractions or width is studied is referred to as the *probe jet*, while the other is referred to as the *reference jet*. The average transverse momentum of the probe and the reference jet is defined as:

$$p_{\mathrm{T}}^{\mathrm{avg}} = \frac{p_{\mathrm{T}}^{\mathrm{probe}} + p_{\mathrm{T}}^{\mathrm{ref}}}{2}. \tag{5.11}$$

Choosing which of the two back-to-back jets is the reference jet and which is the probe jet is arbitrary and the events are always used twice, inverting the roles of reference and probe.

The GS corrections are measured through the asymmetry variable defined as:

$$A(x) = \frac{p_{\mathrm{T}}^{\mathrm{probe}}(x) - p_{\mathrm{T}}^{\mathrm{ref}}}{p_{T}^{\mathrm{avg}}(x)}, \tag{5.12}$$

where x is any of the properties used in the GS calibration (see Table 5.1). Both $p_{\mathrm{T}}^{\mathrm{probe}}$ and $p_{\mathrm{T}}^{\mathrm{ref}}$ depend on x, but the dependence is explicitly written in Equation 5.12 only for the probe jet because the jet property used to build the correction is the one of the probe jet. $p_{\mathrm{T}}^{\mathrm{probe}}$ and $p_{\mathrm{T}}^{\mathrm{ref}}$ are defined with the same calibration. When computing correction number i, they are both corrected up to the $(i-1)^{\mathrm{th}}$ correction (see Section 5.2). The mean response as a function of x is given by:

$$\langle R(x) \rangle = \langle \frac{p_{\mathrm{T}}^{\mathrm{probe}}}{p_{\mathrm{T}}^{\mathrm{ref}}} \rangle, \tag{5.13}$$

which correspond to a good approximation to:

$$\langle R(x) \rangle = \langle \frac{p_{\mathrm{T}}^{\mathrm{probe}}}{p_{\mathrm{T}}^{\mathrm{ref}}} \rangle \simeq \frac{1 + \langle A(x) \rangle / 2}{1 - \langle A(x) \rangle / 2}. \tag{5.14}$$

It was decided to use Equation 5.14 instead of Equation 5.13 to calculate the mean response since $A(x)$ is gaussian in bins of fixed $p_{\mathrm{T}}^{\mathrm{avg}}$. While, the response distribution $R(x)$ calculated as $p_{\mathrm{T}}^{\mathrm{probe}}/p_{\mathrm{T}}^{\mathrm{ref}}$ in bins of $p_{\mathrm{T}}^{\mathrm{ref}}$ is not gaussian distributed.

The measurement of the response through the asymmetry defined in Equation 5.12 assumes that the truth jet level asymmetry is zero. This is true on average but not when computed in bins of x. The measured asymmetry $A(x)$ is therefore a mixture of detector effects and imbalance at the truth level. In order to remove the effect of imbalance at the truth level, a new asymmetry is defined :

$$A'(x) = A(x) - A_{\mathrm{truth}}(x), \tag{5.15}$$

where $A(x)$ is given by Equation 5.12 and $A_{\mathrm{truth}}(x)$ is :

$$A_{\text{truth}}(x) = \frac{p^{\text{probe}}_{\text{T,truth}}(x) - p^{\text{ref}}_{\text{T,truth}}}{p^{\text{avg}}_{\text{T,truth}}(x)}, \tag{5.16}$$

where $p^{\text{avg}}_{\text{T,truth}}(x) = (p^{\text{probe}}_{\text{T,truth}}(x) + p^{\text{ref}}_{\text{T,truth}})/2$. The variable $A_{\text{truth}}(x)$ is the asymmetry at the truth jet level (or truth asymmetry) and is calculated by matching reconstructed jets to truth jets, as explained in Section 5.2.1. A_{truth} can of course only be determined in the Monte Carlo simulation. When using $A'(x)$ instead of $A(x)$ in Equation 5.14, the effects of imbalance at the truth level are removed and the resulting response depends only on detector effects. Accounting for the truth jet imbalance is particularly important for the corrections that depend on the energy in the presampler and the width. This will be discussed in a later section. A more detailed description of the di-jet balance method in the absence of balance at the truth jet level can be found in Ref. [124].

To enhance events with a di-jet topology we use only events where the two leading jets have a $p_{\text{T}}^{\text{avg}} > 20$ GeV and $\Delta\phi > 2.8$ radian, where $\Delta\phi$ is the azimuthal angle between the two leading jets in the event. The study is performed in five representative η bins described in Table 5.5 and in eleven $p_{\text{T}}^{\text{avg}}$ bins (GeV): 20-30, 30-40, 40-60, 60-80, 80-110, 110-160, 160-210, 210-260, 260-310, 310-400, 400-800. The η binning reflects the geometry of the ATLAS calorimeters and guarantee to have enough statistics in each bin to reduce the statistic fluctuations. The trigger selection follows that described in Section 5.3.3. The requirements were chosen such that the trigger efficiency, for each $p_{\text{T}}^{\text{avg}}$ range, was greater than 99% and approximately flat as a function of the pseudorapidity.

$\lvert\eta\rvert$ region	ATLAS detector region
$\lvert\eta\rvert < 0.6$ $0.6 \leq \lvert\eta\rvert < 1.2$	Calorimeter barrel
$1.2 \leq \lvert\eta\rvert < 1.7$ $1.7 \leq \lvert\eta\rvert < 2.1$	Transition region between the barrel and the end-cap
$2.1 \leq \lvert\eta\rvert < 2.8$	End-cap calorimeters

Table 5.5: ATLAS detector regions and pseudorapidity bins used in the analysis.

5.5.2 Validation of the di-jet balance method in the Monte Carlo simulation

Two different validations of the di-jet balance method have been carried out. The first one was done in the reference Pythia sample and compares the response calculated using equation 5.14 to the response calculated using the truth jet as in Equation 5.2. Figure 5.11 shows this comparison for jets after the EM+JES calibration with $80 < p_{\text{T}}^{\text{jet}} < 110$ GeV and $|\eta| < 0.6$. The results obtained using the asymmetry as in Equation 5.12 and when incorporating the truth asymmetry are both shown. If the truth asymmetry is ignored, the calculated response is off by up to 4% for high values of the width and presampler fraction in this particular $p_{\text{T}}^{\text{jet}}$ bin. This difference increases with decreasing $p_{\text{T}}^{\text{jet}}$ reaching 8% for jets of $p_{\text{T}}^{\text{jet}} \approx 20$ GeV. These differences are comparable to the effect to be measured, and reduced to less than 2% when a correction for A_{truth} is used. Similar results have been found in the other $p_{\text{T}}^{\text{jet}}$ and $|\eta|$ bins. After the various GS corrections are applied the jet property dependence observed in Figure 5.11 disappears and the jet response becomes flat and equal to 1 within 1-1.5%.

The second test of validation corresponds to the comparison of the truth asymmetry between different simulated samples. As explained in Section 5.5.1, the truth asymmetry is subtracted from the reconstructed asymmetry in order to remove the imbalance at the truth jet level. Figure

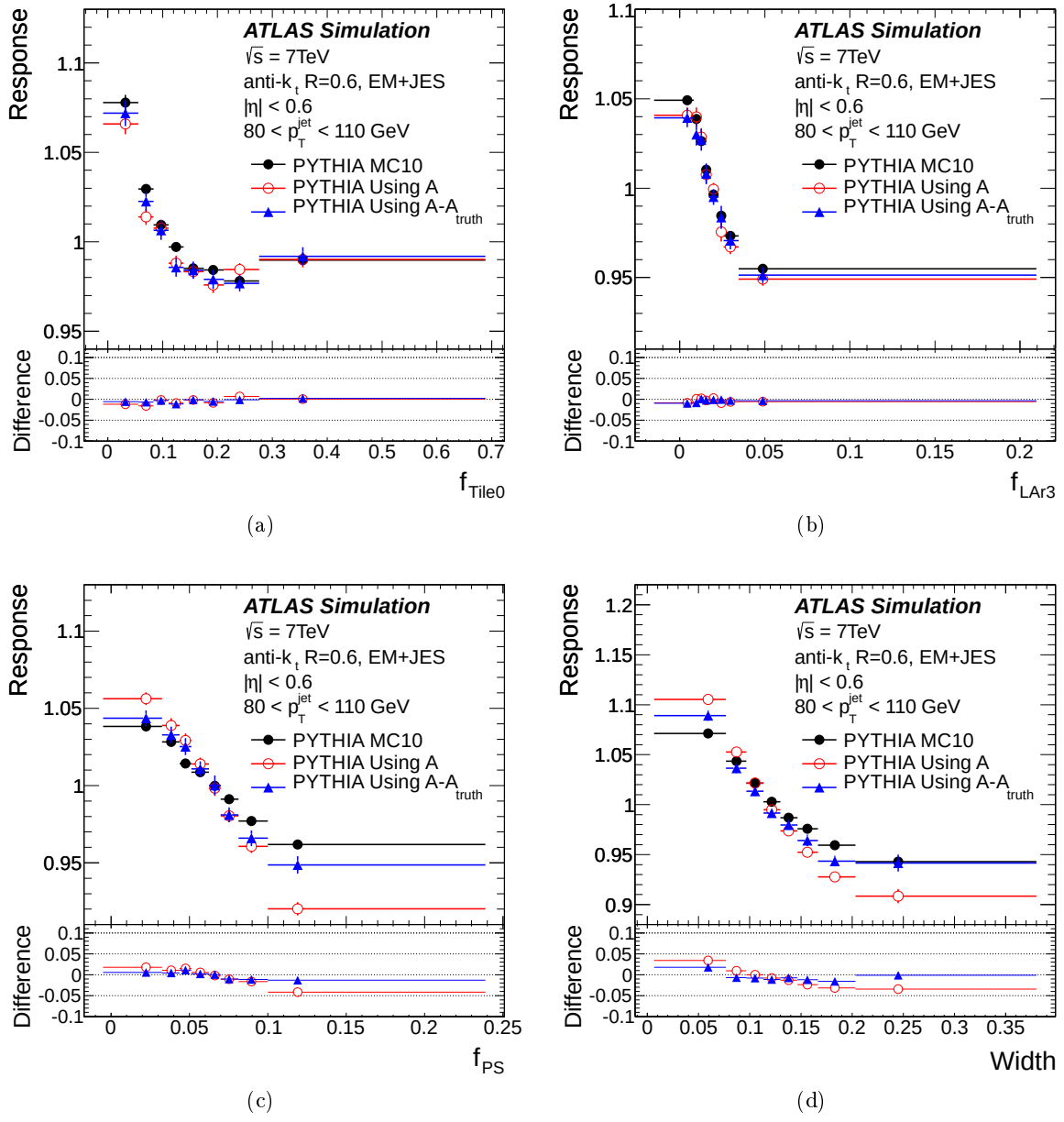


Figure 5.11: Average jet response calculated using truth jets as reference (see Equation 5.2) (full circles), A (open circles), and $A - A_{\text{truth}}$ (triangles) as a function of (a) $f_{\texttt{Tile0}}$, (b) $f_{\texttt{LAr3}}$, (c) f_{PS} and (d) the jet width in the reference PYTHIA sample. The lower part of each figure shows the differences between the response calculated using the truth jet and the one calculated with the di-jet balance method without A_{truth} (blue triangles) and with A_{truth} (open circles). Anti-k_t jets with $R = 0.6$ calibrated with the EM+JES calibrated with $80 < p_{\text{T}}^{\text{jet}} < 110$ GeV and $|\eta| < 0.6$ are used [2].

5.12 shows the truth asymmetry as a function of $f_{\texttt{Tile0}}$, $f_{\texttt{LAr3}}$, f_{PS} and the jet width in the central region for $40 < p_{\text{T}}^{\text{jet}} < 60$ GeV for the various simulated samples described in Section 5.3.2: the reference PYTHIA sample with the MC10 tune, a PYTHIA sample with Perugia2010 tune and a HERWIG++ sample. These last two samples test the sensitivity to the description of soft physics or the specifics of the hadronization process that could cause differences in the truth jet

imbalance. The truth asymmetry differs by no more than 5% in this particular bin. For $p_{\mathrm{T}}^{\mathrm{jet}} > 60$ GeV and other $|\eta|$ bins, the truth asymmetries differ by less than 2%. At low $p_{\mathrm{T}}^{\mathrm{jet}}$ (below 40 GeV in the barrel), the $\Delta\phi$ cut combined with the lower statistics of the Perugia2010 and HERWIG++ samples yield large statistical uncertainties (of the order of 5%), and this statement cannot be made with such precision.

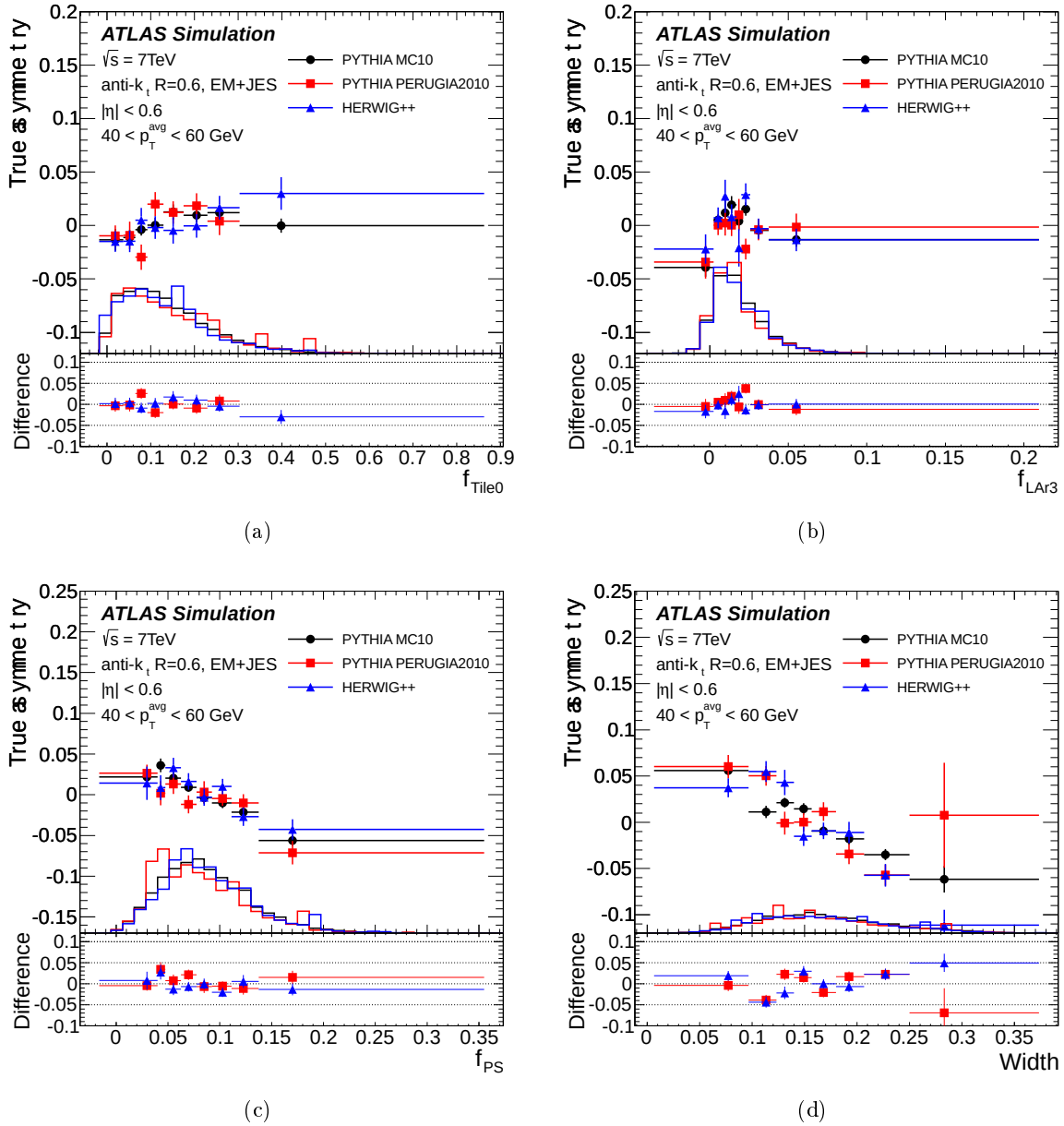


Figure 5.12: Average asymmetry at the truth jet level for different generator tunes (PYTHIA MC10, PYTHIA Perugia2010, HERWIG++) as a function of (a) $f_{\mathtt{Tile0}}$, (b) $f_{\mathtt{LAr3}}$, (c) f_{PS} and (d) the jet width of the probe jet. The jets used have $40 < p_{\mathrm{T}}^{\mathrm{jet}} < 60$ GeV and $|\eta| < 0.6$. The distributions of the jet properties are also shown. The lower part of each figure shows the differences between PYTHIA MC10 and the other Monte Carlo Generators [2].

In summary, the di-jet balance method allows the determination of the response as a function of the layer fractions and the jet width over the entire transverse momentum and rapidity ranges.

It can therefore be used to validate the corrections derived in the Monte Carlo simulation using data.

5.5.3 Differences between data-based and Monte Carlo-based GS corrections

Figure 5.13 shows the difference between the reconstructed asymmetry and the truth asymmetry as a function of $f_{\texttt{Tile0}}$, $f_{\texttt{LAr3}}$, f_{PS} and the jet width for jets at the EM+JES scale with $80 < p_{\mathrm{T}}^{\mathrm{jet}} < 110$ GeV and $|\eta| < 0.6$. The truth asymmetry used is the one from the reference Pythia sample. The disagreement between data and the reference Pythia sample in this bin is 2.5% at worse. Similar results are found in the others bins for the layer fractions. For the width, disagreements of up to 2% are found in other regions except for $1.7 < |\eta| < 2.1$, where disagreements of up to 5% are observed for low values of the jet width.

The asymmetries in data shown in Figure 5.13, and the corresponding asymmetries in other $p_{\mathrm{T}}^{\mathrm{jet}}$ and $|\eta|$ bins, are used to derive data-based corrections. Figure 5.14 compares the average jet response calculated using the truth jet (using Equation 5.2) as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ after the Monte Carlo-based and data-based corrections are applied to the reference Pythia sample. The uncertainties are bigger when applying data-based corrections than when applying MC-based corrections. Given that the sample used is the same (the reference Pythia sample), the difference comes from the fact that the uncertainty on data-based corrections are bigger. The calculation of the uncertainties on data-based corrections is described in Appendix C. Figure 5.13 provides a quantitative measure of how differences between the data and Monte Carlo simulation impact the jet energy scale and could be used (in combination with results in Section 5.4.2) in the future to establish a systematic uncertainty for GS due to changes in the topology and flavor composition. After the first two corrections in Table 5.1 the response changes by less than 1% for data-based and Monte Carlo-based corrections. The response changes by an additional 1% to 2% after the third (f_{PS}) and the fourth (width) corrections are applied in the barrel. The agreement in the end-cap is within 2% (4%) for $p_{\mathrm{T}}^{\mathrm{truth}} > 60$ GeV (< 60 GeV).

The jet transverse momentum resolution calculated using the truth jet (using Equation 5.2) as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ after the Monte Carlo-based and data-based corrections are applied to the reference Pythia sample is shown in Figure 5.15 for the different detector regions. The resolution improvement when using the data-based corrections are comparable with the one achieved when using MC-based corrections for $|\eta| < 1.7$. For $|\eta| > 1.7$ the improvement is smaller. The difference is due to the difference in the width distribution between data and simulation as will be discussed in next section.

Figure 5.16 shows the average jet response after applying data-based GS corrections to the simulation in bins of width 0.1 in $|\eta|$ for $80 < p_{\mathrm{T}}^{\mathrm{truth}} < 110$ GeV and $260 < p_{\mathrm{T}}^{\mathrm{truth}} < 310$ GeV. As explained in Section 5.5.1 (see Table 5.5), data-based corrections have been determined in wider $|\eta|$ regions than Monte Carlo-based corrections. The application of data-based corrections therefore introduces a variation of the response with $|\eta|$, as can be seen from Figure 5.16. They reflect the limitations of the di-jet balancing method given the statistics available at the time this study was performed.

Data-based corrections have been also derived with truth asymmetries coming from the Pythia Perugia2010 and Herwig++ samples shown in Figure 5.12. These corrections are then applied to the reference Pythia MC10 sample and the response they yield is compared to the response obtained after applying the reference data-based corrections using the truth asymmetry from the reference Pythia MC10 sample. The difference in response is found to be lower than 1% in all the $p_{\mathrm{T}}^{\mathrm{jet}}$ and $|\eta|$ bins where the statistical uncertainty is sufficient to derive such a number (except for $1.2 < |\eta| < 1.7$ and 60 GeV$< p_{\mathrm{T}} <$80 GeV where the difference is found to be around 4%) as can be seen in Figure 5.17 for $|\eta| < 0.6$ and $1.2 < \eta < 1.7$.

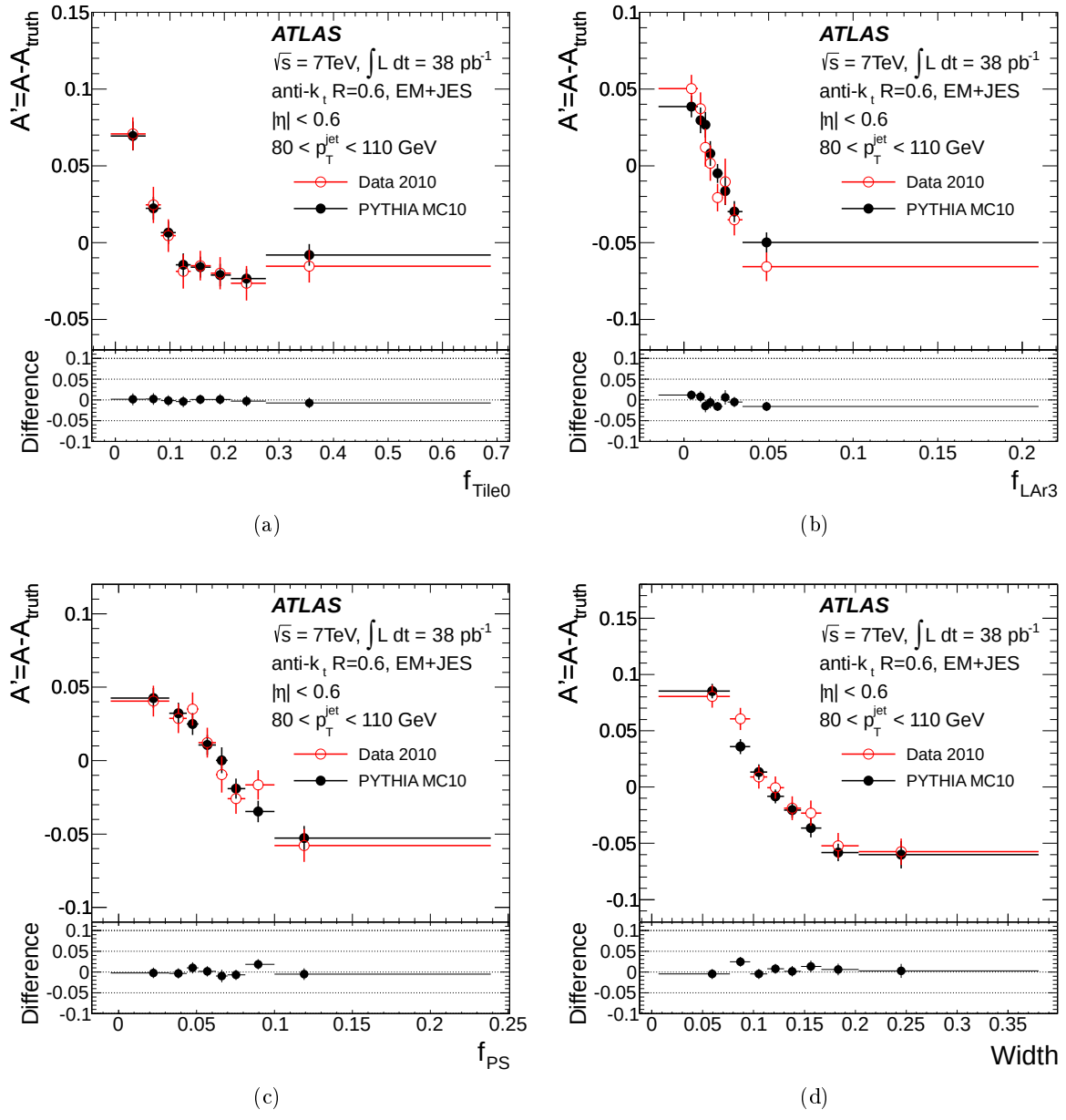
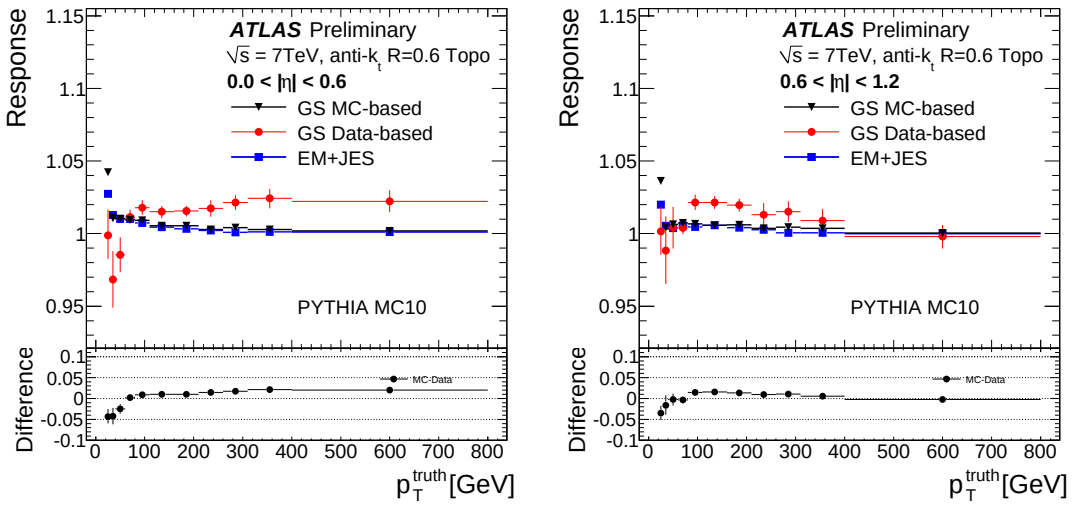
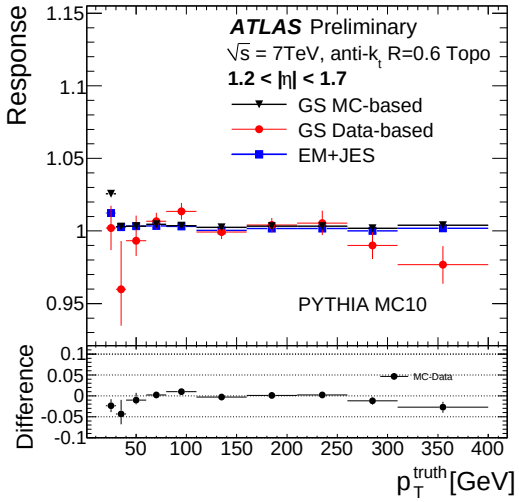


Figure 5.13: Difference between the average reconstructed asymmetry and the truth asymmetry in data (open circles) and in the reference PYTHIA sample (full circles) as a function of (a) $f_{\texttt{Tile0}}$, (b) $f_{\texttt{LAr3}}$, (c) f_{PS} and (d) the jet width of the probe jet. The lower part of each figure shows the differences between data and Monte Carlo simulation. The jets used have $80 < p_{\mathrm{T}}^{\mathrm{jet}} < 110$ GeV and $|\eta| < 0.6$ and are calibrated with the EM+JES scheme.

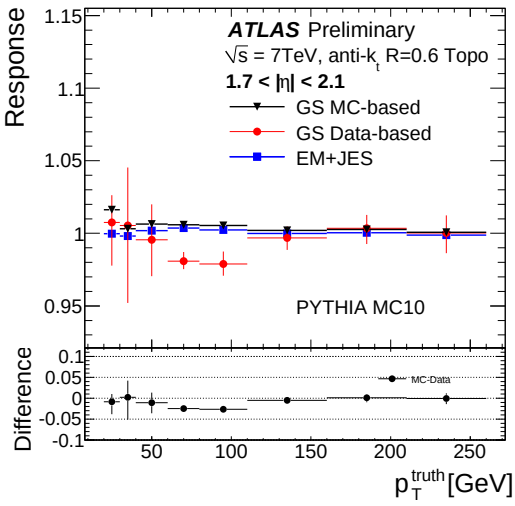


(a)

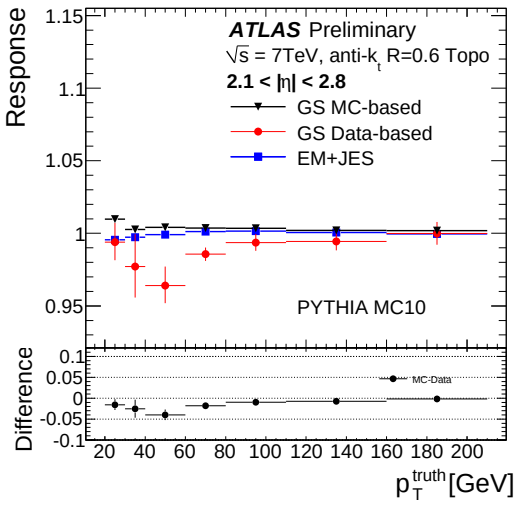
(b)



(c)

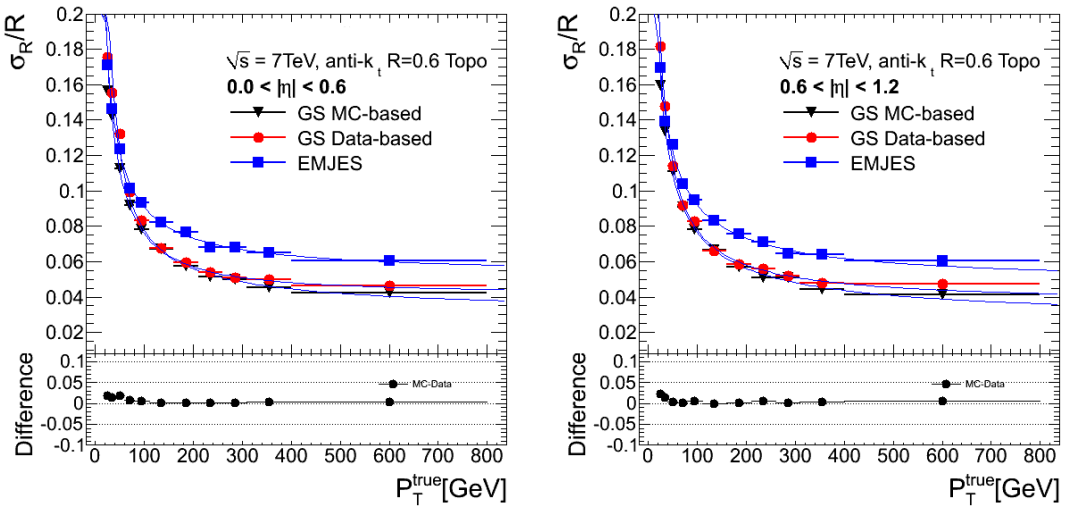


(d)



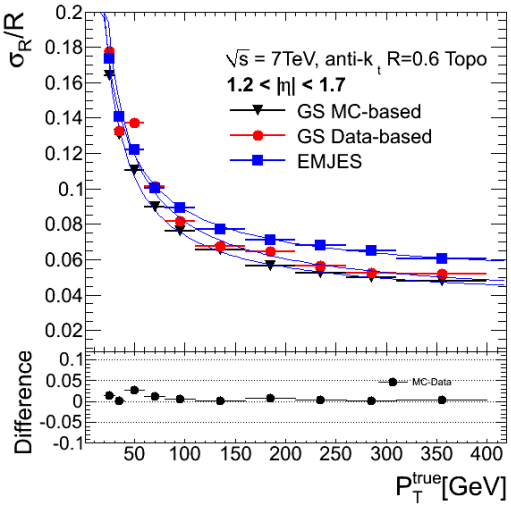
(e)

Figure 5.14: Average jet response calculated using the truth jet as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ after applying corrections derived from the nominal PYTHIA sample (black triangles) and data (red circles) for jets with (a) $|\eta| < 0.6$, (b) $0.6 < |\eta| < 1.2$, (c) $1.2 < |\eta| < 1.7$, (d) $1.7 < |\eta| < 2.1$ and (e) $2.1 < |\eta| < 2.8$. Differences between data and Monte Carlo simulation responses are shown at the bottom of each figure.

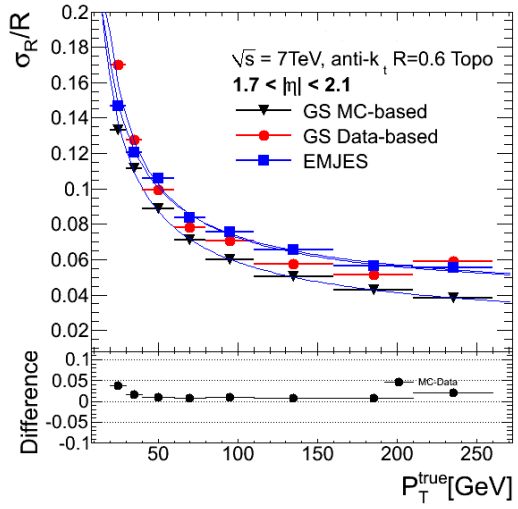


(a)

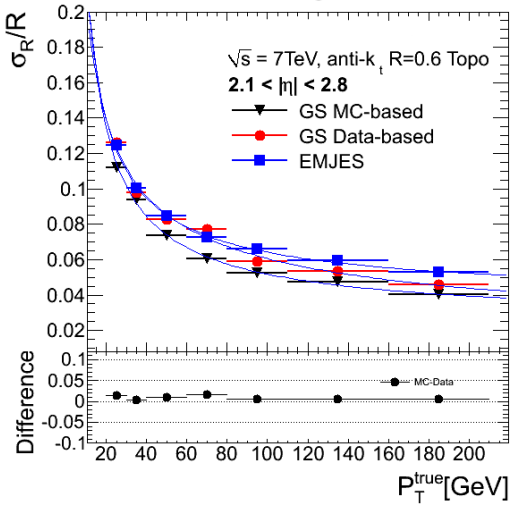
(b)



(c)

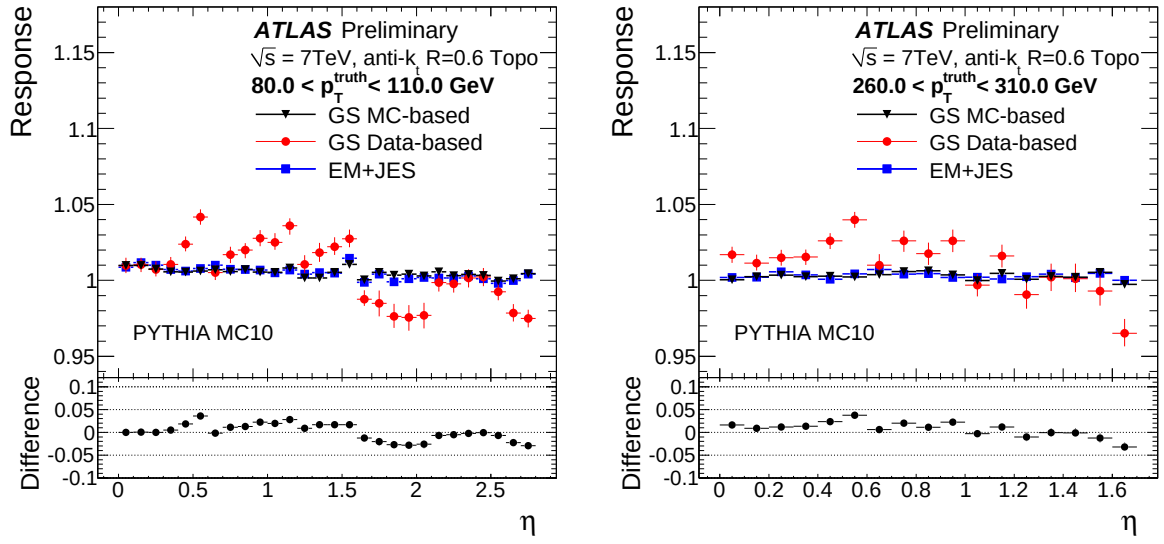


(d)



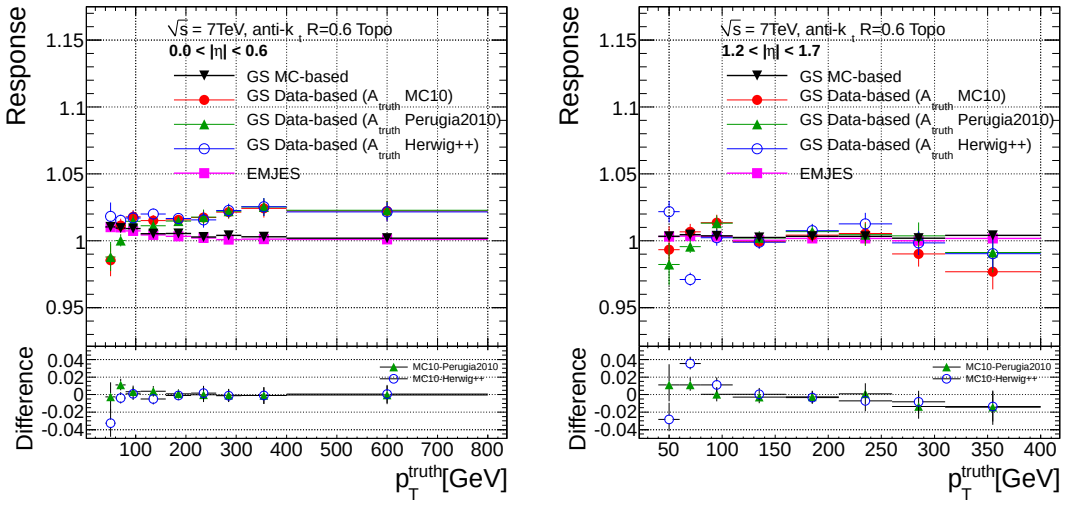
(e)

Figure 5.15: Jet transverse momentum resolution calculated using the truth jet as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ after applying corrections derived from the nominal Pythia sample (black triangles) and data (red circles) for jets with (a) $|\eta| < 0.6$, (b) $0.6 < |\eta| < 1.2$, (c) $1.2 < |\eta| < 1.7$, (d) $1.7 < |\eta| < 2.1$ and (e) $2.1 < |\eta| < 2.8$. Differences between data and Monte Carlo simulation responses are shown at the bottom of each figure.



(a) (b)

Figure 5.16: Average jet response calculated using the truth jet as a function of η after applying the GS corrections derived from the nominal PYTHIA sample (black triangles) and data (red circles) for jets with (a) $80 < p_{\mathrm{T}}^{\mathrm{truth}} < 110$ GeV and (b) $260 < p_{\mathrm{T}}^{\mathrm{truth}} < 310$ GeV. Differences between data and Monte Carlo simulation responses are shown at the bottom of each figure.



(a) (b)

Figure 5.17: Average jet response calculated using the truth jet as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ after applying GS corrections derived from the nominal PYTHIA sample (black triangles) and data using the A_{truth} calculated using the nominal PYTHIA MC10 (red circles), the Perugia2010 tune (green triangles) and the HERWIG++ sample (blue circles) for jets with (a) $|\eta| < 0.6$ and (b) $1.2 < |\eta| < 1.7$.

5.6 Evaluation of the systematic uncertainty associated to the GS calibration

5.6.1 Evaluation of the systematic uncertainty using inclusive multi-jets events

In Section 5.5, data-based GS corrections were derived and compared to Monte-Carlo based corrections by applying both corrections to the same sample (the reference PYTHIA sample). In this section, the approach is different: the same GS corrections (here the Monte Carlo-based ones) are applied to both data and nominal Monte Carlo inclusive samples. The difference between data and simulation therefore reflects differences in the jet properties used as input to the GS calibration in the inclusive samples. Differences in the jet properties used as input to GS between data and Monte Carlo simulation may lead to a degradation of linearity when Monte Carlo-based corrections are applied to data. For instance, if the $f_{\texttt{Tile0}}$ property is lower on average in data than in Monte Carlo simulation in some $p_{\mathrm{T}}^{\mathrm{jet}}$ and η bin, the coefficient applied to jets will be lower in data than in Monte Carlo simulation, since the response decreases with $f_{\texttt{Tile0}}$, as can be seen from Figure 5.11. The response will therefore be lower in data than in simulation.

Figure 5.18 shows the mean value of f_{PS}, $f_{\texttt{LAr3}}$, $f_{\texttt{Tile0}}$ and width as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ in the barrel for data and various Monte Carlo samples: the reference PYTHIA MC10, PYTHIA Perugia2010 and HERWIG++. The agreement for $f_{\texttt{Tile0}}$ and f_{PS} between data and PYTHIA MC10 is within 5% over the entire $p_{\mathrm{T}}^{\mathrm{jet}}$ range. For $f_{\texttt{LAr3}}$, this agreement is also within 5% except for $20 < p_{\mathrm{T}}^{\mathrm{jet}} < 30$ GeV where a disagreement of 7.5% is observed. A larger disagreement is found for the jet width. Jets are 5% (10%) wider in data than in Monte Carlo simulation at 200 GeV (600 GeV). Figure 5.19 shows the standard deviation of the properties as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ in the barrel. The agreement between data and PYTHIA MC10 for $f_{\texttt{LAr3}}$ and f_{PS} is within 5% over the entire $p_{\mathrm{T}}^{\mathrm{jet}}$ range. For $f_{\texttt{Tile0}}$ and the jet width, disagreements of 10% are observed in some $p_{\mathrm{T}}^{\mathrm{jet}}$ bins. Similar results are found in the other $|\eta|$ bins for the layer fractions. For the jet width, the disagreement between data and Monte Carlo simulation is slightly worse in the range $2.1 < |\eta| < 2.8$ than in the other ranges.

Figures 5.18 and 5.19 show that the reference PYTHIA MC10 and the PYTHIA Perugia2010 tunes agree to within a few percent. The agreement of the HERWIG++ sample with data is found to be as good as for the other samples for $f_{\texttt{LAr3}}$ and $f_{\texttt{Tile0}}$, except for $20 < p_{\mathrm{T}}^{\mathrm{jet}} < 30$ GeV. For f_{PS} and the jet width, disagreements of 5 to 10% are found between HERWIG++ and the other samples for $p_{\mathrm{T}}^{\mathrm{jet}} < 60$ GeV. For $p_{\mathrm{T}}^{\mathrm{jet}} > 160$ GeV, HERWIG++ is found to describe better the width observed in data than the other samples.

The systematic uncertainty can be quantitatively estimated by comparing the difference in the correction coefficients $E_{\mathrm{GS}}^{\mathrm{jet}}/E_{\mathrm{EM+JES}}^{\mathrm{jet}}$ between data and Monte Carlo simulations. Figure 5.20 shows the correction coefficient as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ in the barrel in data and in the reference PYTHIA sample after GSL and GS. Figure 5.21 shows the same quantity but as a function of η for $80 < p_{\mathrm{T}}^{\mathrm{jet}} < 110$ GeV. Deviations from the unity in the ratios between data and Monte Carlo simulation as shown in the lower part of Figures 5.20 and 5.21 represent the systematic uncertainty associated to the GS corrections. This uncertainty is added in quadrature to the EM+JES uncertainty [3, 116].

The results for all the $p_{\mathrm{T}}^{\mathrm{jet}}$ and η ranges are the following: for $20 < p_{\mathrm{T}}^{\mathrm{jet}} < 30$ GeV and $|\eta| < 2.1$, the uncertainty varies from 0.5% to 0.7% depending on the $|\eta|$ region. For $p_{\mathrm{T}}^{\mathrm{jet}} > 30$ GeV and $|\eta| < 2.1$, the uncertainty is lower than 0.5%. For $2.1 < |\eta| < 2.8$, the uncertainty varies from 0.4% to 1% depending on the $p_{\mathrm{T}}^{\mathrm{jet}}$ bin. For a given $p_{\mathrm{T}}^{\mathrm{jet}}$, the uncertainty is higher for $2.1 < |\eta| < 2.8$ than for $|\eta| < 2.1$ because of the poorer description of the jet width. For $2.1 < |\eta| < 2.8$ the GSL scheme shows slightly larger difference than the GS scheme. In general, the uncertainty is

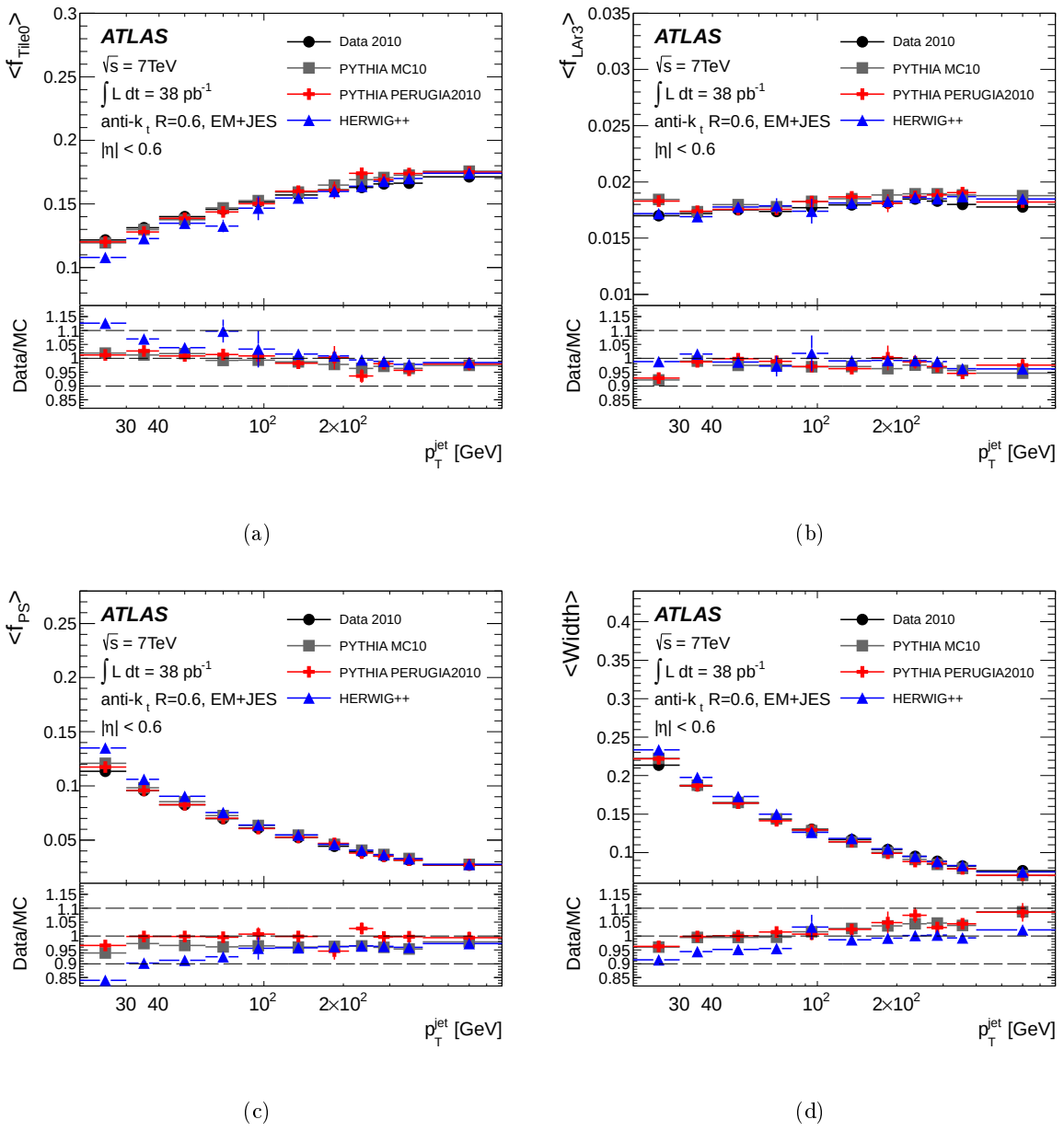


Figure 5.18: Mean value of (a) f_{PS}, (b) $f_{\mathtt{LAr3}}$, (c) $f_{\mathtt{Tile0}}$ and (d) the jet width as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ for $|\eta| < 0.6$ for data and various Monte Carlo simulations. Jets are calibrated with the EM+JES calibration. The ratio of data over Monte Carlo responses are shown at the bottom.

lower than 1% for $20 < p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV and $0 < |\eta| < 2.8$.

It is worth noting that the uncertainty coming from the imperfect description of the jet properties described in this section and the differences between data-based and Monte Carlo-based corrections presented in Section 5.5.3 are not independent. Indeed, the average response after GS in each $p_{\mathrm{T}}^{\mathrm{jet}}$ and $|\eta|$ bin, which depends on both the distribution of the properties and the GS corrections, is constrained to be close to the response after the EM+JES calibration. A change in the distribution of a jet property therefore translates into a change in the GS correction as a function of this property such that the average response remains unchanged. The differences described in Section 5.5.3 are therefore partly caused by differences in the jet properties.

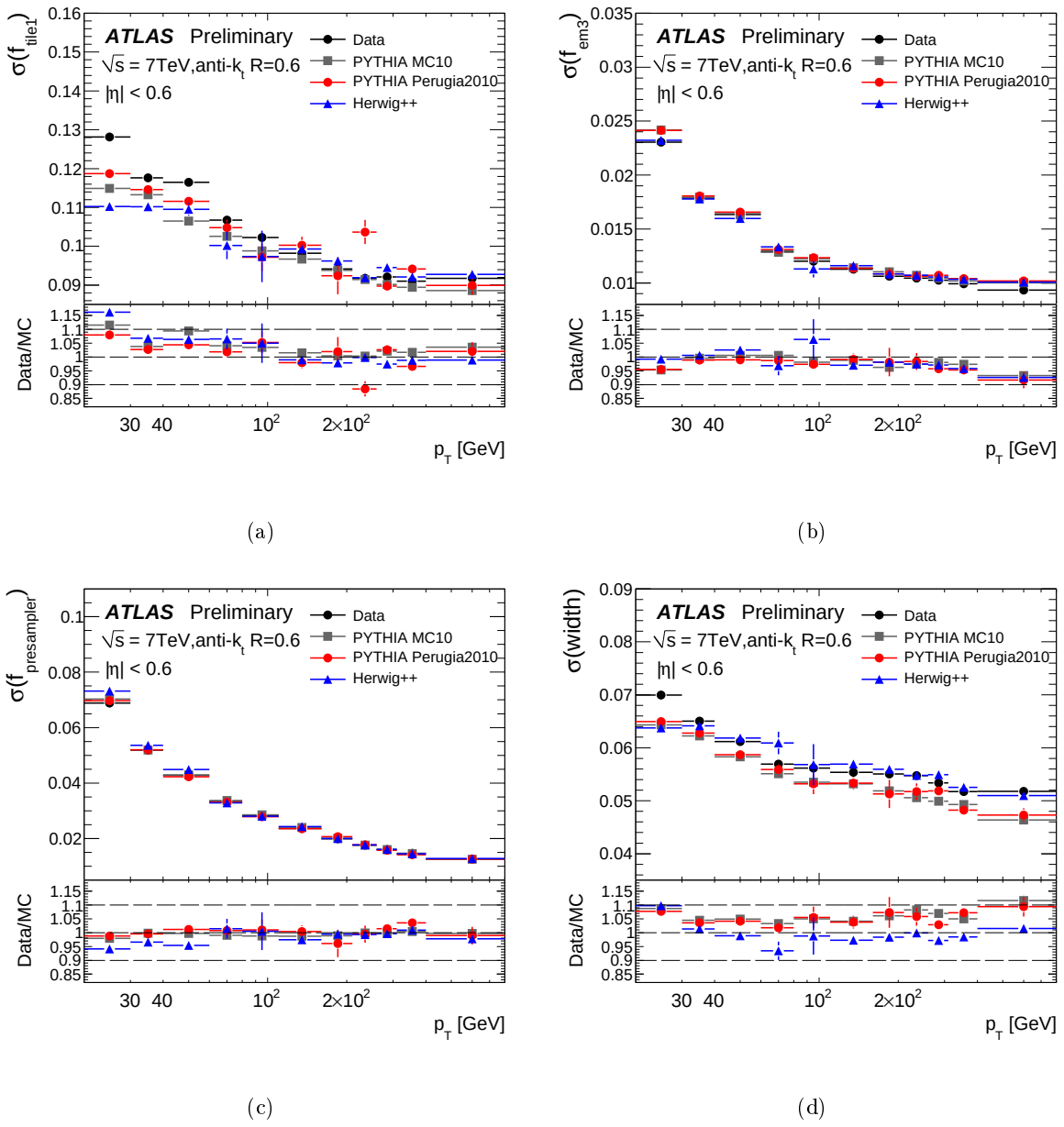
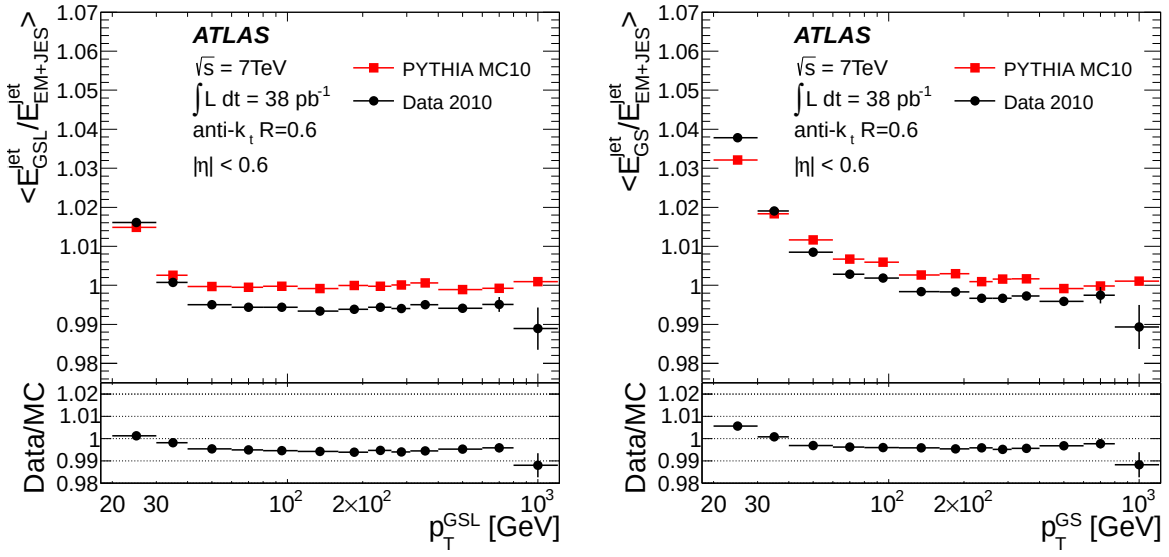


Figure 5.19: Standard deviation of (a) f_{PS}, (b) $f_{\mathtt{LAr3}}$, (c) $f_{\mathtt{Tile0}}$ and (d) the jet width as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ for $|\eta| < 0.6$ for data and various Monte Carlo simulations. Jets are calibrated with the EM+JES calibration. The ratio of data over Monte Carlo responses are shown at the bottom.

5.6.2 Evaluation of the systematic uncertainty using γ+jet events

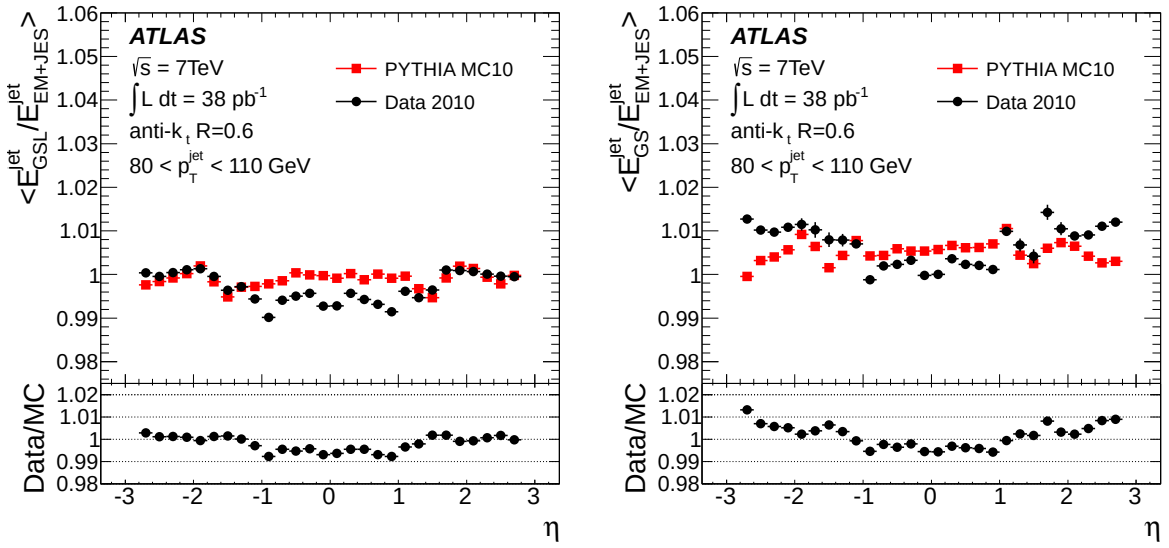
The jet energy scale after each correction in the GS calibration can also be verified using other common data-based jet energy scale validation methods. One of them, discussed in this section, is the p_{T} balance method in γ+jet events.

Two methods based on γ+jet events have been described in detail in [133]. The method that balances directly a photon against a jet is used because it is simpler to apply to jets of different calibrations and because the other method (MPF) would require a missing transverse energy calculation that accounts for the changes in the jet energy scale introduced by the GS



(a) (b)

Figure 5.20: Average jet energy after (a) GSL and (b) GS divided by the average jet energy after the EM+JES calibration as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ in data and the reference PYTHIA sample in the barrel. The double ratio $(E_{\mathrm{GS(L)}}/E_{\mathrm{EM+JES}})_{\mathrm{data}}/(E_{\mathrm{GS(L)}}/E_{\mathrm{EM+JES}})_{\mathrm{MC}}$ is shown at the bottom.



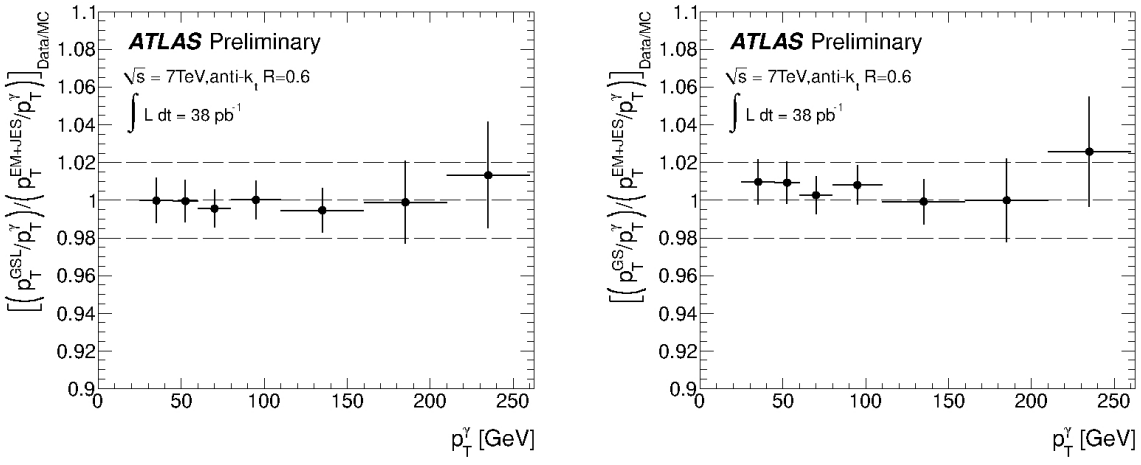
(a) (b)

Figure 5.21: Average jet energy after (a) GSL and (b) GS divided by the average jet energy after the EM+JES calibration as a function of η in data and the reference PYTHIA sample for $80 < p_{\mathrm{T}}^{\mathrm{jet}} < 110$ GeV. The double ratio $(E_{\mathrm{GS(L)}}/E_{\mathrm{EM+JES}})_{\mathrm{data}}/(E_{\mathrm{GS(L)}}/E_{\mathrm{EM+JES}})_{\mathrm{MC}}$ is shown at the bottom.

calibration coherently. The measurement is repeated with jets calibrated with the EM+JES calibration and after the application of each of the corrections that build up the GS calibration. The datasets, event selection and analysis cuts used are the same as in [133]. The response is defined as the ratio between the jet and the photon transverse momenta. Only one η bin is used ($|\eta| < 1.2$) to maximize the available statistics. The approach to evaluate the systematic

uncertainty is similar to that described in Section 5.6.1: the Monte Carlo-based GS corrections are applied to both data and Monte Carlo simulation and the systematic uncertainty associated to GS is evaluated by computing the data/Monte Carlo ratio of the response after GS relative to EM+JES.

For $25 < p_{\mathrm{T}}^{\mathrm{jet}} < 45$ GeV, the agreement between the response in data and Monte Carlo simulation is 3.2% after EM+JES and 4.2% after GS. For $210 < p_{\mathrm{T}}^{\mathrm{jet}} < 260$ GeV, the agreement is 5% after EM+JES and 2.5% after GS. The systematic uncertainty is shown in figure 5.22. The errors bars represent statistical uncertainties. The systematic uncertainty varies from 1% at $p_{\mathrm{T}}^{\mathrm{jet}} = 25$ GeV to 2.5% for $p_{\mathrm{T}}^{\mathrm{jet}} = 260$ GeV. These results are compatible within statistical uncertainty with the uncertainty evaluated using inclusive multi-jets (see Section 5.6.1).



(a) (b)

Figure 5.22: Ratio of the detector response in data and the nominal PYTHIA sample using direct p_T balance in γ+jets events after (a) GSL and (b) GS relative to EM+JES. The jets used have $|\eta| < 1.2$.

5.6.3 Final systematic uncertainty for the GS calibration

Data-driven methods for estimating the systematic uncertainty in events with different topology and flavour composition (di-jet and γ+jet events) have also been presented, but the statistic available at the time this study was performed was not enough to quantify the uncertainty in such samples with high precision over the entire p_{T} and η range. Therefore, it was decided to quote the systematic uncertainty calculated in the inclusive multi-jets sample as the final systematic uncertainty on the GS calibration. It was found to be lower than 1% for $|\eta| < 2.8$ and $20 < p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV.

This uncertainty has to be added in quadrature to the EM+JES calibration systematic uncertainty. The addition in quadrature was motivated by the fact that the systematic uncertainties for EM+JES and GS can be considered as independent. The possible correlations between the sources of the EM+JES systematic uncertainty [116, 3] and the jet properties are briefly discussed below:

- Dead material could be responsible at least partly for the imperfect description of the jet properties, but its contribution to the EM+JES systematic uncertainty is small.
- The calorimeter cell noise threshold can change the cluster shapes and the fake cluster multiplicity, affecting the jet properties distribution. But this effect is only important for

20 GeV< p_{T} <30 GeV and is of the order of 2%. Added linearly to the GS systematic uncertainty in that bin would lead to a systematic uncertainty of around 2.7%, which is not big.

- The EM+JES non closure reflects the self consistency of the EM+JES calibration procedure. Therefore, it does not depend on the jet properties.

- The calorimeter response is mainly intrinsic to the detector and should not depend on the jet physics.

- The EM+JES systematic uncertainty due to physics model and parameters employed in the Monte Carlo generator was evaluated by comparing the nominal Monte Carlo samples with others generators and tunes: Alpgen and Perugia2010. The differences in the jet properties distributions between the Monte Carlo nominal sample and Perugia2010 were studied and found to be in most bins smaller or equal to the differences between the nominal samples and data (see Figure 5.18). In addition, its contribution to the EM+JES systematic uncertainty is of the order of 1-2%, which is bigger that the systematic uncertainty for GS. It looks like the contribution to the EM+JES systematic is coming mainly from another effect than differences in the GS jet properties distributions.

The GS corrections were derived in a Monte Carlo simulation samples with a particular flavour composition (mixture of light quarks, heavy quarks and gluon initiated jets) and with a given selection of isolated jets. Therefore, the GS systematic uncertainty calculated in the inclusive multi-jets sample are only valid under these particular conditions. Extra systematic uncertainties need to be added to account for the dependence of the jet response on the jet topology and flavour. Initial studies that can be used in the future (using more statistics) were presented in Sections 5.5.3 and 5.6.2.

5.7 GS calibration and other jet calibration schemes

Due to the non-compensating nature of the ATLAS calorimeter, signal losses due to noise thresholds and in dead material the jet energy needs to be calibrated. Four different calibration schemes have been developed in ATLAS. Taking into account the performance and data validation studies presented along this chapter (see Sections 5.4 and 5.5) and similar studies performed using the other jet calibration schemes [3], a qualitative comparison between the four calibration schemes developed can be done based on five important criteria:

Jet resolution Results presented in Section 5.4.1 shown that the jet energy resolution achieved in data and simulation by GCW, LCW and GS are very similar. In real data the improvement with respect to EM+JES is found to be up to 30 % at 400 GeV for GS, LCW and GCW. The worst jet resolution corresponds to the EM+JES calibration.

Systematic uncertainty derivation The EM+JES is the simplest calibration among the four calibration schemes and the derivation of its systematic uncertainty is the easiest one. In terms of complexity it is followed by the GS systematic uncertainty derived in this chapter. These uncertainties are also supported by in situ techniques. The JES uncertainties in the LCW+JES and GCW+JES jet calibration schemes can be only derived from a combination of several in situ techniques [3]. At the time of these studies, they suffer from the limited number of events in data samples.

Flavour sensitivity After the results discussed in Section 5.4.2 the different jet calibration schemes can be ordered according to the flavor sensitivity criteria in increasing order as: GS, GCW, LCW and EM+JES.

Derivation of the calibration from data The EM+JES, GS and GCW are global calibrations in the sense that the jet is calibrated as a whole. Therefore, these calibrations can be derived from data using data-driven techniques. This is not the case for the LCW calibration, where the corrections are applied at the topo-cluster level, before reconstructing the jet.

Subjet calibration Jets produced in very high energy collisions with a high boost are close to each others and can be reconstructed into one fat jet. These fat jets have an internal structure. This internal structure can be easily calibrated using the LCW calibration since the jet constituents are calibrated, without considering the jet context. The GCW calibration can be used too without major complications as weights used to calibrate the jets depend on the calorimeter cell energy density. This task would be more complicated for calibrations as the EM+JES and GS, where the jet constituents are at the EM scale.

$E_{\mathrm{T}}^{\mathrm{miss}}$ **calibration** As described in Section 4.9 the $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed from the sum of the topological clusters. LCW is the most suitable calibration to use for the $E_{\mathrm{T}}^{\mathrm{miss}}$, since it treats the effects of non-compensation, dead material deposits and out-of-cluster deposits at the topocluster level. EM+JES and GS are situated in the other extreme. They calibrate the jet as a whole, without taking into account cell or topoclusters individual information. GCW can be considered as being between LCW and EM+JES/GS. GCW also calibrate jets as a whole, but the weights used in the correction depend on the calorimeter cell energy density.

Chapter 6

Search for $t\bar{t}$ resonances in ATLAS

Contents

6.1 Overview

This chapter describes the search of new heavy particles decaying into top-quark pairs using the ATLAS detector at the LHC. Several beyond-the-SM theories predict the existence of such heavy resonances (see Section 1.3 in Chapter 1), with a very low production cross section compared to the Standard Model $t\bar{t}$ pair production. The $t\bar{t}$ decay topology considered in this analysis corresponds to the lepton plus jets topology (see Section 1.4 in Chapter 1). Its branching fraction is smaller than the one for the full hadronic topology. However, it is experimentally favored since the presence of a lepton in the final state allows to reduce the multijet background contamination in the analysis. On the other hand, the dileptonic topology provides a clear signature as two leptons are found in the final state, but the branching fraction is small and

the full reconstruction of the top pair presents some ambiguities. The ambiguity arises from the presence of two neutrinos in the final state. The presence of an additional neutrino in the dileptonic topology prevents the association between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the neutrino's p_{T}.

Only the "e plus jets" and "μ plus jets" final state topologies are considered. They result from the decay chain in which one of the W boson from the top-quark decay decays into an electron and an electron neutrino or into a muon and a muon neutrino, respectively, and the other W boson decays into a quark-antiquark pair. The "τ plus jets" final state topology is not directly considered, since taus are difficult to identify, as discussed already in Chapter 4. However, taus decaying leptonically can be misidentified as electrons or muons. taus decaying hadronically can be misidentified as jets or even as electrons when they produce narrow jets with low track multiplicity. Therefore a fraction of "τ plus jets" events can be selected in the analysis. The reconstructed $t\bar{t}$ mass spectrum is the discriminant variable used in the search. In principle, the top pair is reconstructed from six individually identified decay products: four jets, an electron or a muon and a neutrino (in the form of $E_{\mathrm{T}}^{\mathrm{miss}}$). However, at the LHC energies top quarks can be produced with a large transverse momentum compared to its mass. They are known as boosted tops. Boosted tops are expected to be produced for resonance masses above approximately 1 TeV [134]. Usually the boosted top decay products can not be resolved individually since they are very close to each other. The event selection used in this analysis has been optimized to include events where two of the hadronic decay products were possibly reconstructed as a single jet. A highly boosted top topology correspond to one where the three top decay products are reconstructed as a single object. A search for highly boosted top pair resonances in ATLAS is described in Ref. [135].

The beyond-the-SM models used as benchmarks in this search are discussed in Section 6.2. A brief summary of the results from the previous searches for $t\bar{t}$ resonances is presented in Section 6.3. Data and Monte Carlo samples used in the analysis are described in Section 6.4. The object reconstruction and selection is described in Section 6.5. Followed in Section 6.6 by a discussion of the event selection used and the performance of the different methods used to reconstruct the $t\bar{t}$ pair in Section 6.9. The determination of the different backgrounds is described in Section 6.7. In Section 6.10 the different systematic uncertainties affecting the study are listed. Finally, the results are presented in Section 6.11.

6.2 Theoretical benchmarks

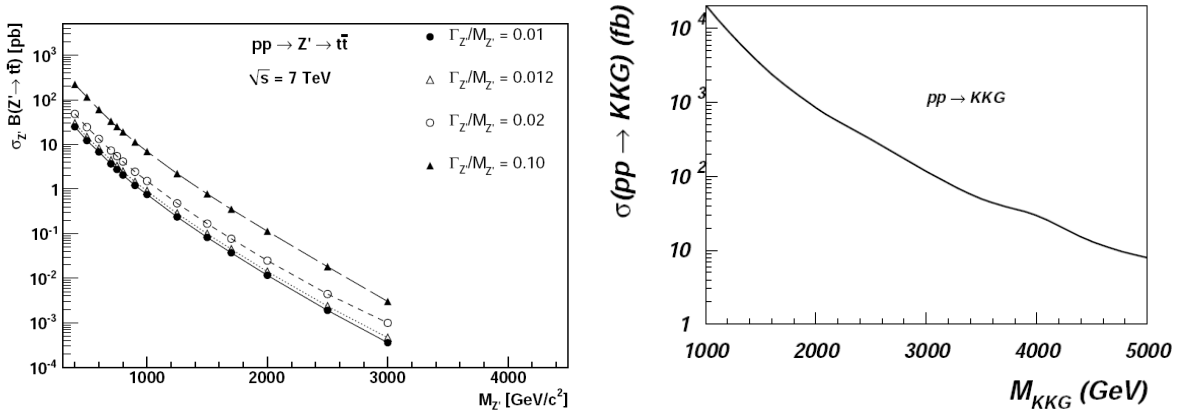
Many models of beyond-the-SM physics predict the existence of resonances decaying into top-quark pairs (see Chapter 1 Section 1.3). In this analysis two benchmark scenarios are used:

- A topcolor Z' boson [30] arising in Topcolor Assisted Technicolor (see Section 1.3.2). In the scenario considered in this analysis the Z' boson is not generated by imbedding $U(1)$, but by using instead a leptophobic interaction. In the leptophobic scenario considered here[1], this resonance couples strongly only to the first and third generation of quarks and has not significant couplings to leptons, which enhances the branching ratio of the Z' decay into $t\bar{t}$ pairs. It is said to be leptophobic and topophyllic. This specific model is used to allow direct comparison with DØ and CDF Collaboration's results [136, 137] and because it gives a significant cross section at the LHC at $\sqrt{s} = 7$ TeV as can be seen in Figure 6.1a. The Z' decaying into $t\bar{t}$ pairs produces a narrow peak in the $t\bar{t}$ invariant mass spectrum. Therefore, it is the detector resolution who dominates the width of the reconstructed peak.

[1] Specifically, model **IV** in Ref. [30] with $f_1 = 1$ (to enhance the coupling to the $t\bar{t}$ channel) and $f_2 = 0$ (to avoid the coupling to the $b\bar{b}$ channel) and a width of 1.2% of the Z' boson mass is used.

- A Kaluza-Klein gluon g_{KK}, which appears in Randall-Sundrum models with warped extra dimensions described in Section 1.3.3. The model used is a color octet resonance with spin 1, with no couplings to leptons (and will go unobserved in the dilepton channel). The couplings to quarks take the RS values used in [34]: $g_L = g_R = -0.2g_S$ for light quarks including charm, $g_L = 1.0g_S$, $g_R = -0.2g_S$ for bottom quarks and $g_L = 1.0g_S$, $g_R = 4.0g_S$ for the top quark. In this chosen scheme, the resonance is predicted to be significantly wider than the detector resolution. The branching fraction $BR(g_{KK} \to t\bar{t})$ is estimated to be 92.5% and $BR(g_{KK} \to b\bar{b}) = 5.5\%$ (the remaining 2% corresponds to light quark jets). The g_{KK} cross section at the LHC (with its nominal conditions) is shown for illustratives purpose in Figure 6.1b. This model should be considered as a proxy for colored resonances, since precision constraints force the mass of the resonance for this specific choice of couplings to be above about 2 TeV.

Cross-sections for the different models for various resonance or threshold masses are given in Section 6.4.2.2.



(a) (b)

Figure 6.1: The total cross section of (a) Z' at $\sqrt{s} = 7$ TeV [30] and (b) KK gluon at $\sqrt{s} = 14$ TeV [35] production at the LHC as a function of its respective mass.

6.3 Existing limits

Previous searches for $t\bar{t}$ resonances were most recently carried out by the CDF [138, 139, 140, 137, 141] and DØ [136, 142] collaborations at Run II of the Fermilab Tevatron Collider, and by the ATLAS [143, 7, 135] and CMS [144, 145, 146, 147, 148] experiments at the LHC. At Tevatron experiments no evidence for new particles was found and 95% limits were set on the mass of a leptophobic topcolor Z' boson at $m_{Z'} > 900$ GeV as well as on the coupling strength of a heavy colour-octet vector particle. A similar scenario was found at the LHC, limits were set on the mass of a leptophobic topcolor Z' and a KK gluon. The best limit obtained for the Z' with a width of 1.2% of the Z' boson mass and the KK gluon in the lepton plus jets resolved topology has been obtained by the CMS experiment, $500 < m_{Z'} < 1300$ and $1000 < m_{g_{KK}} < 1400$, respectively, using 4.7 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV collected in 2011 [145]. The $t\bar{t}$ invariant mass was used as the final discriminant variable and the limits were set using a frequentist CLs approach. A summary of the different searches results can be found in Table 6.1.

Channel	$\mathcal{L}$	Z' 1.2% (GeV)	Z' 10%(GeV)	g_{KK} (GeV)
CDF experiment				
Lepton+jets [137]	4.8 fb^{-1}	$450 < m_{Z'} < 900$		
Full hadronic [141]	2.8 fb^{-1}	$450 < m_{Z'} < 805$		
DØ experiment				
Lepton+jets [142]	5.3 fb^{-1}	$350 < m_{Z'} < 835$		
ATLAS experiment				
Dilepton [143]	1.04 fb^{-1}			$500 < m_{g_{KK}} < 840$
Lepton+jets [7]	200 pb^{-1}	$\sigma_{Z'} < 38$ (3.2) pb for Z' 0.5 (1) TeV		$500 < m_{g_{KK}} < 650$
CMS experiment				
Dilepton [144]	5.0 fb^{-1}	$500 < m_{Z'} < 1100$		
Lepton+jets resolved [145]	4.7 fb^{-1}	$500 < m_{Z'} < 1300$	$500 < m_{Z'} < 1700$	$1000 < m_{g_{KK}} < 1400$
e/μ+jets boosted [146, 147]	4.3 fb^{-1}/1.1 fb^{-1}	$\sigma_{Z'} < 2.51$ (0.62) pb for Z' 1 (2) TeV		
Full hadronic [148]	4.6 fb^{-1}	$1300 < m_{Z'} < 1500$	$1000 < m_{Z'} < 2000$	$1400 < m_{g_{KK}} < 1500$

Table 6.1: Summary of the results of Tevatron and LHC $t\bar{t}$ resonance searches. Results correspond to a 95% of confidence level. Results for two Z' resonances are presented with a width of 1.2% and 10% of the Z' boson mass.

6.4 Data and event selection for 2011 data

6.4.1 Data sample

The data sample used corresponds to pp collisions at $\sqrt{s} = 7$ TeV collected by the ATLAS detector between March and August 2011, using a single muon or single electron trigger. It corresponds to data periods between B and K processed with the release 16 of the ATLAS software. The trigger applied varied with data periods and can be found in Table 6.2. Only data where all relevant ATLAS subdetectors were operational was used. The data sample corresponds to an integrated luminosity of approximately 2.05 ± 0.08 fb^{-1} [56, 57].

Data period	$\mathcal{L}$ (pb^{-1})	Electron channel	Muon channel
B-I	1340.3	`EF_e20_medium`	`EF_mu18`
J	212.2	`EF_e20_medium`	`EF_mu18_medium`
K	500.0	`EF_e22_medium`	`EF_mu18_medium`

Table 6.2: Triggers used in the different data periods between March and August 2011. The number in the trigger name denote the approximate value of the trigger threshold on the lepton p_{T}.

6.4.2 Monte Carlo simulation

The signal and most relevant background samples have been generated at $\sqrt{s} = 7$ TeV. After event generation all simulated samples were run through the standard GEANT4 [91, 92] simulation of the ATLAS detector and reconstructed and analyzed using the same software used for data. The ATLAS detector geometry used in the simulation of the nominal sample reflects the geometry of the detector as best known at the time of these studies. Simulated events were required to pass the simulated `EF_mu18` and `EF_e20_medium` in the muon and electron channel respectively.

6.4.2.1 Background processes

There are a certain number of processes that can emulate a final state similar to the one generated by our benchmark signals decaying into $t\bar{t}$ pairs in the lepton plus jets channel. The following SM background processes were simulated using the Monte Carlo technique:

- The SM $t\bar{t}$ background was simulated using the MC@NLO v3.41 generator with the CTEQ6.6 PDF set [149], and showered using HERWIG v6.5 in association with JIMMY. During the generation process the top-quark mass was set to 172.5 GeV. Only events in which at least one of the W bosons decays leptonically were produced, corresponding to a cross-section of 79.99 pb to which a K-factor of 1.117 was applied to account for NNLO corrections [150].

- The electroweak single top-quark production was simulated using the same programs used to generate the SM $t\bar{t}$ background. Only leptonic W boson decays were required for the s- and t-channel processes. For the Wt process, all decays were produced. The inclusive cross-sections are based on approximate NNLO calculations: 64.57 pb (t-channel) [46], 4.63 pb (s-channel) [47] and 15.74 pb (Wt process) [48].

- W and Z plus jets samples with leptonic vector boson decays were simulated with the ALPGEN v2.13 generator in exclusive bins of parton multiplicity for multiplicities lower than five, and inclusively above that. The CTEQ6L1 [151] PDF set was used. $Wc\bar{c}$, Wc and $Wb\bar{b}$ events were generated separately with ALPGEN and double counting was avoided by removing events with $b-$ and $c-$quarks from the $W+$light jet samples based on a ΔR matching. The $Z+$jets events include Z events and Drell-Yan contribution from the $\gamma* \rightarrow \ell\ell$. The $Z/\gamma*$ interference was taken into account. $Z+$jets events were required to have a dilepton invariant mass $40 < m_{\ell\ell} < 2000$ GeV. The events were showered with HERWIG and JIMMY and matching was performed with the MLM method. The MLM method removes overlaps between the n and $n+1$ parton samples. The cross-sections for the different subsamples are given in Tables 6.3, 6.4 and 6.5. Flat K-factors equal to 1.20 and 1.25 were applied on the W and $Z+$jets samples respectively based on NNLO QCD calculations [152]. In addition, the $W+$jets samples were normalized to data as described in Section 6.7.2.

- Diboson samples have been produced using HERWIG v6.5 with MRST2007LO* [153] PDFs with JIMMY. A filter requiring at least one lepton with $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.8$ was applied. The cross-sections (K-factors) used for these filtered samples are: 11.05 pb (1.48) for WW production, 3.46 pb (1.60) for WZ production and 0.97 pb (1.30) for ZZ production. The K-factors used are such that the cross-sections agree with results obtained using the MCFM generator [152]. They were calculated based on NLO corrections.

- Others $t\bar{t}$ samples have been produced with different configurations to evaluate the systematic uncertainty associated to the modeling of the shape of the $t\bar{t}$ mass distribution

due to the MC generator used (comparing the nominal SM $t\bar{t}$ sample against a sample generated with POWHEG+HERWIG), the parton shower model (POWHEG+HERWIG samples compared to POWHEG+PYTHIA samples) and ISR and FSR effects (by comparing the nominal SM $t\bar{t}$ sample against samples produced using ACERMC+PYTHIA using different settings for the ISR and FSR modelization[2]).

Subsample	Cross section (pb)
$W \to e\nu$ + 0lp	6921.60
$W \to e\nu$ + 1lp	1304.30
$W \to e\nu$ + 2lp	378.29
$W \to e\nu$ + 3lp	101.43
$W \to e\nu$ + 4lp	25.87
$W \to e\nu$ + 5lp	7.00
$W \to \mu\nu$ + 0lp	6919.60
$W \to \mu\nu$ + 1lp	1304.20
$W \to \mu\nu$ + 2lp	377.83
$W \to \mu\nu$ + 3lp	101.88
$W \to \mu\nu$ + 4lp	25.75
$W \to \mu\nu$ + 5lp	6.92
$W \to \tau\nu$ + 0lp	6918.60
$W \to \tau\nu$ + 1lp	1303.20
$W \to \tau\nu$ + 2lp	378.18
$W \to \tau\nu$ + 3lp	101.51
$W \to \tau\nu$ + 4lp	25.64
$W \to \tau\nu$ + 5lp	7.04

Table 6.3: Cross-sections for the various $W \to \ell\nu$+jets subsamples. K-factors equal to 1.20 are applied.

Subsample	Cross section (pb)
$W + b\bar{b}$ + 0lp	47.32
$W + b\bar{b}$ + 1lp	35.77
$W + b\bar{b}$ + 2lp	17.34
$W + b\bar{b}$ + 3lp	6.63
$W + c\bar{c}$ + 0lp	127.53
$W + c\bar{c}$ + 1lp	104.68
$W + c\bar{c}$ + 2lp	52.08
$W + c\bar{c}$ + 3lp	16.96
$W + c$ + 0lp	644.4
$W + c$ + 1lp	205.0
$W + c$ + 2lp	50.8
$W + c$ + 3lp	11.4
$W + c$ + 4lp	2.8

Table 6.4: Cross-sections for the various W+heavy flavour jets subsamples. The W boson decays into a lepton (e, μ or τ) plus the corresponding ν. K-factors equal to 1.20 are applied.

6.4.2.2 Signal Processes

Samples for topcolor Z' bosons were generated with PYTHIA v6.421, allowing all three $t\bar{t}$ topologies. A K-factor of 1.3 was applied to Z' samples to account for NLO corrections [30]. Samples for KK gluons were generated with MADGRAPH v4.4.51 showered with PYTHIA. The KK gluon samples cross-sections were recalculated using PYTHIA v8.1 (LO). The CTEQ6L1 PDF set was used in the generation of the signal samples. The effect of interference with the SM processes production has not been taken into account as the backgrounds and signal samples were produced separately. The cross-sections for the signals subsamples can be found in Table 6.6.

[2]The parameters controlling the ISR/FSR emission in PYTHIA are PARP(67) and PARP(64) for ISR and PARP(72) and PARJ(82) for FSR. The nominal values correspond to PARP(67)=4.0, PARP(64)=1.0, PARP(72)=0.192 GeV and PARJ(82)=1.0 GeV. To decrease (increase) the ISR emission PARP(67) and PARP(64) are set to 0.5 and 4.0 (6.0 and 0.25), respectively. To decrease (increase) the FSR emission PARP(72) and PARJ(82) are set to 0.096 GeV and 2.0 GeV (0.384 GeV and 0.5 GeV), respectively. Six samples with different settings were generated: ISR increased, ISR decreased, SFR increased, SFR decreased and a simultaneous increasing and decreasing of ISR and FSR.

Subsample	Cross-section (pb)
$Z \to ee$ + 0lp	668.32
$Z \to ee$ + 1lp	134.36
$Z \to ee$ + 2lp	40.54
$Z \to ee$ + 3lp	11.16
$Z \to ee$ + 4lp	2.88
$Z \to ee$ + 5lp	0.83
$Z \to \mu\mu$ + 0lp	668.68
$Z \to \mu\mu$ + 1lp	134.14
$Z \to \mu\mu$ + 2lp	40.33
$Z \to \mu\mu$ + 3lp	11.19
$Z \to \mu\mu$ + 4lp	2.75
$Z \to \mu\mu$ + 5lp	0.77
$Z \to \tau\tau$ + 0lp	668.40
$Z \to \tau\tau$ + 1lp	134.81
$Z \to \tau\tau$ + 2lp	40.36
$Z \to \tau\tau$ + 3lp	11.25
$Z \to \tau\tau$ + 4lp	2.79
$Z \to \tau\tau$ + 5lp	0.77

Table 6.5: Cross-sections for the various $Z \to \ell\ell$+jets subsamples. K-factors equal to 1.25 are applied.

6.5 Object reconstruction and selection

The reconstruction of the different objects in the final state of events with a $t\bar{t}$ topology was described in Chapter 4. In this analysis physics objects are required to satisfy the following requirements:

Electrons has to be of quality Tight. They must also have $E_{\mathrm{T}} > 15$ GeV and $0 < |\eta_{\mathrm{cluster}}| < 2.47$. Electrons in the calorimeter transition region $1.37 < |\eta_{\mathrm{cluster}}| < 1.52$ are excluded. They are also required to be isolated: `EtCone20` < 3.5 GeV. Additionally, electron candidates falling within the LAr hole are rejected.

Muons are required to be tight combined Muid muons with $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$. The isolation requirement is `etcone30`< 4.0 GeV and `ptcone30`< 4.0 GeV. Any muon candidate within $\Delta R = 0.4$ of a jet with $p_{\mathrm{T}} > 20$ GeV is rejected, to avoid contamination from muons produced in a jet, for example from B hadrons.

Jets are reconstructed from topoclusters with the anti-k_t algorithm with radius parameter $R = 0.4$. They are calibrated at the EM+JES scale. Jets are required to have a $p_{\mathrm{T}} > 20$ GeV and $\eta < 4.5$. The closest jet within $\Delta R < 0.2$ from a selected electron is removed, to avoid double counting of electrons as jets.

b−tagged jets are tagged using the `JetFitterCombNN` algorithm with an operating point of 2.00, which corresponds to a $b-$tagging efficiency of 60% and a light quark jet rejection of 345 in simulated $t\bar{t}$ events.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ used is the refined $E_{\mathrm{T}}^{\mathrm{miss}}$. It is reconstructed as indicated in Chapter 4 Section 4.9, where the electrons were defined of the tight type to be consistent.

Sample	Cross-section × BR($Z'/g_{KK} \to t\bar{t}$) (pb)
Topcolor Z'	NLO
$m_{Z'_t} = 500$ GeV	19.6
$m_{Z'_t} = 600$ GeV	10.3
$m_{Z'_t} = 700$ GeV	5.6
$m_{Z'_t} = 800$ GeV	3.2
$m_{Z'_t} = 900$ GeV	1.9
$m_{Z'_t} = 1000$ GeV	1.2
$m_{Z'_t} = 1200$ GeV	0.46
$m_{Z'_t} = 1400$ GeV	0.19
$m_{Z'_t} = 1600$ GeV	0.086
$m_{Z'_t} = 1800$ GeV	0.039
$m_{Z'_t} = 2000$ GeV	0.018
g_{KK}	LO
$m_{g_{KK}} = 500$ GeV	81.2
$m_{g_{KK}} = 600$ GeV	39.4
$m_{g_{KK}} = 700$ GeV	20.8
$m_{g_{KK}} = 800$ GeV	11.6
$m_{g_{KK}} = 900$ GeV	6.8
$m_{g_{KK}} = 1000$ GeV	4.1
$m_{g_{KK}} = 1200$ GeV	1.7
$m_{g_{KK}} = 1400$ GeV	0.73
$m_{g_{KK}} = 1600$ GeV	0.35
$m_{g_{KK}} = 1800$ GeV	0.18
$m_{g_{KK}} = 2000$ GeV	0.095

Table 6.6: Cross-sections for the resonant signal processes.

The corresponding scaling factors and energy corrections described along Chapter 4 are applied to the reconstructed objects. The uncertainties on them are used to determined the corresponding systematic uncertainties.

6.6 Event selection

From Section 6.4.2 it is evident that our signal cross-sections are smaller than the background processes cross-sections. First a set of cuts are applied to clean-up the samples. Then, extra selection cuts are applied in order to enhance the signal over background fraction. The event selection used in this analysis is the following:

Clean-up selection

1. Events are required to pass the Good Run List selection defined by the ATLAS Top Group. This first cut concerns only data. More information about the Good Run List can be found in Section 4.1.

2. The event is required to be accepted by the electron or muon trigger (see Section 6.4).

3. A primary vertex reconstructed with at least five tracks originating from it is required. This requirement improves the rejection of non collision background events from the underlying event and/or cosmic radiation.

4. Events with a jet pointing in the direction of the LAr hole are rejected (see Section 4.1), as well as events with noise bursts and/or data integrity errors in the LAr calorimeter. Noise bursts can lead to fake energy deposits in calorimeter cells that can be reconstructed as fake jets or electrons.

5. If an event contains any bad jet [154] with $p_{\mathrm{T}} > 20$ GeV and $E > 0$ GeV the event is rejected. This cut helps to remove non-collision background from noise in the calorimeters, cosmic rays or beam induced background. The criteria used to define a bad jet are almost the same as the one described in Section 5.3.4, except that the jet timing was required to be greater than 25 ns to avoid out-of-time jets coming from cosmic rays for example. This cut is only applied to real data events.

Signal enhancement selection

1. Events were required to contain at least one isolated lepton: one electron with $E_{\mathrm{T}} > 25$ GeV (electron channel) or one muon with $p_{\mathrm{T}} > 25$ GeV (muon channel). The lepton requirement reduces the number of multijets background events, which do not have a real lepton. The electron p_{T} cut is set to 25 GeV since the triggers used in the electron channel reach their efficiency plateau at that value. In the muon channel the trigger efficiency plateau starts at 20 GeV, but the the muon p_{T} cut is set to 25 GeV to reduce the multijet contamination.

2. Events are required not to contain a second lepton with $p_{\mathrm{T}} > 15$ GeV. It mainly helps to reject diboson background events, Z+jets and SM $t\bar{t}$ events decaying in the dileptonic topology.

3. In the electron channel, the selected lepton is required to match the trigger lepton. Due to a problem in data processing, this requirement is not applied in the muon channel. This requirement has a small effect in the selection efficiency. The matching criteria is required since the electron SFs used are valid only for the electrons matched to the electron that fired the lepton trigger.

4. Events where the electron shared an inner detector track with a non isolated muon are rejected.

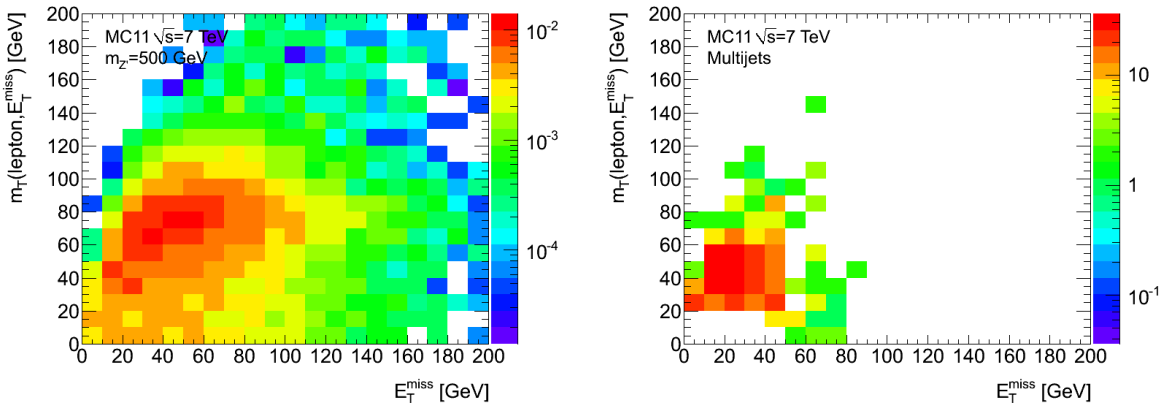
5. In the electron channel, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be larger than 35 GeV and the lepton+$E_{\mathrm{T}}^{\mathrm{miss}}$ transverse mass, M_{T}, larger than 25 GeV. In the muon channel, the requirements are $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}}$+$M_{\mathrm{T}} > 60$ GeV. The $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement helps mainly to reduce the multijet and Z+jets background events, which do not have neutrinos in their final state. Events with a fake lepton as the multijet events have lower $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} than events were a W is produced. Figure 6.2 shows the M_{T} as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$ for the multijet and Z' samples. The cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} is harder in the electron channel than in the muon channel since the first one suffers from a higher multijet background contamination.

6. Events are required to have at least 3 jets with ${p_{\mathrm{T}}}^{\mathrm{jet}} > 25$ GeV and $|\eta^{\mathrm{jet}}| < 2.5$ if one of the jets has mass $m^{\mathrm{jet}} > 60$ GeV. Otherwise, at least four jets with ${p_{\mathrm{T}}}^{\mathrm{jet}} > 25$ GeV and $|\eta^{\mathrm{jet}}| < 2.5$ are required. This cut helps to reduce almost all the backgrounds, except the SM $t\bar{t}$, since they are characterized by a low jet multiplicity. The resolved $t\bar{t}$ topology where the decay products are well separated is characterized by 4 jets in its final state. However,

there is the possibility that one of these jets was not reconstructed due to inefficiencies in the jet reconstruction process or that the $t\bar{t}$ pair is boosted (see discussion in Section 6.1). In the last case, the momenta of the top quark and W boson would be large enough for some of the decays to be merged into one single jet. Previous studies has shown that the jet invariant mass can be used to identify the $t\bar{t}$ topology [134]. The invariant mass of jets coming from a W boson or a top quark is related to the mass of the parent particle. In case of partially merged events the product decays of the W boson are expected to be merged into one fat jet with a mass close to the W mass. It can also happen that the $b-$quark produced in the top-quark decay merges with one of the W boson decay products. In case of fully merged events, all the product decays from the top quark are merged into one fat jet with a mass close to the top-quark mass. The jet invariant mass region between 60 GeV and 150 GeV is characterized by the partially merged topology. Figure 6.3 shows the invariant jet mass distribution for different Z' samples. The number of events with $m^{\text{jet}} > 60$ GeV represents less than 1% of the selected sample. In the following, events with at least one jet with mass $m^{\text{jet}} > 60$ GeV will be referred to as high mass events.

7. At least one of the jets is required to be tagged as a b-jet.

8. The leading p_{T} jet is required to have $p_{\text{T}}{}^{\text{jet}} > 60$ GeV. This requirement reduces the contamination from W/Z+jets and multijet backgrounds, which usually have jets with low p_{T} in their final states.



(a) (b)

Figure 6.2: Missing transverse energy, $E_{\text{T}}^{\text{miss}}$, as a function of the W boson transverse mass, M_{T} for (a) Z' events with $m_{Z'} = 500$ GeV and for (b) multijet events.

The selection efficiencies for the different background and signal samples are presented in Table 6.7 and Figure 6.4. The Z' selection efficiency raises with increasing $m_{Z'}$ until $m_{Z'} = 1$ TeV. After $m_{Z'} = 1$ TeV the selection efficiency decreases due to the lepton isolation requirement, because the lepton and the jet from the W leptonic decay might be close-by. The selection efficiency in the muon channel is higher than in the electron channel. The criteria used in the electron identification, reconstruction and isolation are tighter than those used for the muon, in order to avoid misidentification of jets as electrons. The selection efficiency for the g_{KK} is compatible with the Z' one, given the statistical uncertainties and the larger width of the resonance, which degradates the dependence of the selection efficiency as a function of the resonance mass. SM $t\bar{t}$ events constitute the highest background, followed by the W+jets and the single

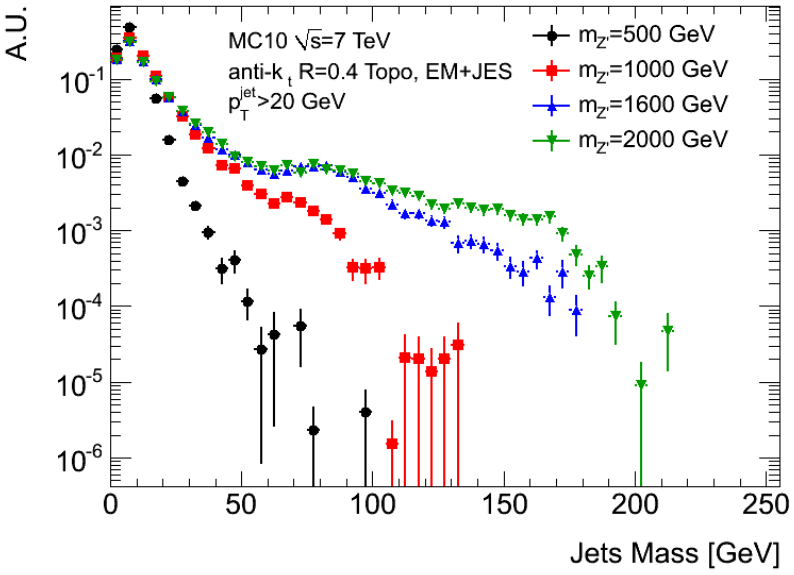


Figure 6.3: Invariant jet mass distribution for Z' samples with different masses. A bump around the W boson mass is observed for the samples with $m_{Z'} = 1000$ GeV, $m_{Z'} = 1600$ GeV and $m_{Z'} = 2000$ GeV. A small bump around the top-quark mass is observed in highest Z' mass sample.

top events. The W+jets selection efficiency is of the order of 0.002%, but its production cross section is some orders of magnitude higher than for the SM $t\bar{t}$. The multijet background will be described in Section 6.7.3. A detailed comparison between data and expectation after the event selection described above is presented in Section 6.8.

Sample	W+jets	Z+jets	$t\bar{t}$	single-top	Diboson
e	0.0017	1.459×10^{-5}	4.270	0.573	0.036
μ	0.0021	1.152×10^{-5}	5.456	0.699	0.035

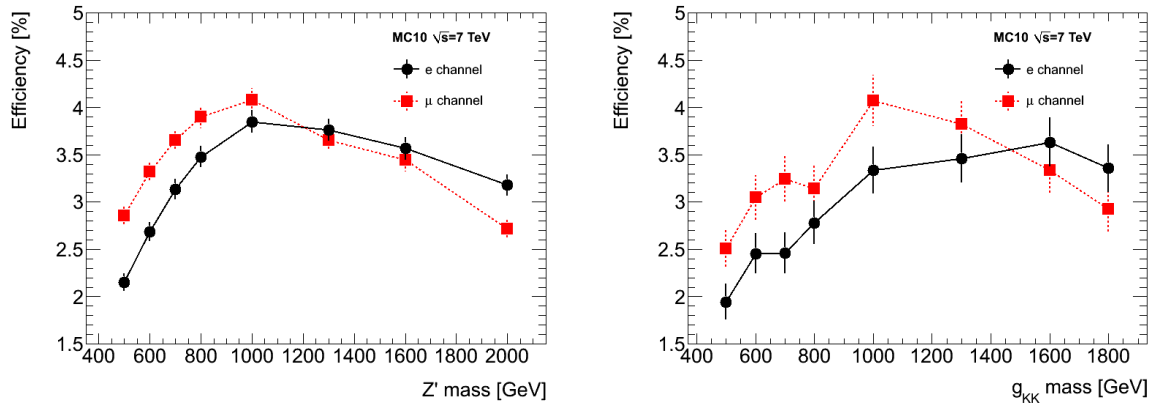
Sample	$m_{KK}500$	$m_{KK}1000$	$m_{KK}1600$	Z' 500	Z' 1000	Z' 1600
e	1.944	3.336	3.630	2.150	3.847	3.564
μ	2.504	4.070	3.333	1.944	3.336	3.630

Table 6.7: Selection efficiency [%] for electroweak backgrounds and some signal samples.

6.7 Background determination

6.7.1 SM $t\bar{t}$ and single top

As described in Section 6.4 the SM $t\bar{t}$ and single top samples were generated using the MC@NLO v3.41 generator with the CTEQ6.6 PDF set. The modeling of the $t\bar{t}$ background is extremely important in this analysis as it constitutes the main background. To evaluate the impact of the PDF choice the events in the SM $t\bar{t}$ and single top samples were reweighted to the MRST2008NLO PDF set (see Section 3.2.2.3). It was found that the agreement between data and expected



(a) (b)

Figure 6.4: Selection efficiency as a function of the signal mass for (a) Z' and (b) g_{KK}. The Z' and g_{KK} samples contain all $t\bar{t}$ decay topologies.

background in the angular distributions was better when using MRST2008NLO. Figures 6.5 and 6.6 show the distributions for the jet and lepton rapidities and azimuthal angles between them and the $E_{\mathrm{T}}^{\mathrm{miss}}$ for both PDF sets after applying the events selection described in Section 6.6. It was decided to reweight the SM $t\bar{t}$ and single top samples to MRST2008NLO. The difference between both PDF sets is used to calculate the systematic uncertainty associated to the PDF choice.

6.7.2 W+jets background

W+jets background is the second largest background in the analysis. It is determined using the ALPGEN samples described in Section 6.6. The normalization of W+jets events can be better estimated by using data control samples than Monte Carlo. Therefore, two additional normalization factors derived using data-driven techniques are applied: W+jets overall normalization and W+jet flavor fractions. The first one attempts to match the number of W+jets in the simulation to the data-driven data yields calculated in off-signals regions. The second one scales each flavor component ($W + b\bar{b}$+jets, $W + c\bar{c}$+jets, $W + c$+jets and W+light jets) to match the flavor fractions calculated in data in off-signal regions. The overall W+jets normalization factors are derived based on the charge asymmetry in W-boson production at the LHC [155]:

$$(N_{W^+} + N_{W^-})^{\mathrm{expected}} = \left(\frac{r_{MC}+1}{r_{MC}-1}\right)(N_{W^+} - N_{W^-})^{\mathrm{data}}, \tag{6.1}$$

where N_{W^+} and N_{W^-} are the number of events with W^+ and W^- bosons and $r_{MC} = N_{W^+}/N_{W^-}$. The difference, $(N_{W^+} - N_{W^-})^{\mathrm{data}}$, and ratio, r_{MC}, are extracted from data and simulation, respectively, as a function of the number of reconstructed jets and $b-$tagged jets. The overall normalization scale factors in the $b-$tagged inclusive 4-jets bin is found to be 0.906 for the electron channel and 0.814 for the muon channel. The overall normalization uncertainty on the W+jets background is set at 48%.

The flavour composition is determined from data based on the tagged fraction of one and 2-jets events [156] and the known $b-$tagging efficiencies (see Section 4.8). A system of two equations expressing the number of W+1 jet ans W+2 jets events before and after $b-$tagging is built. They are expressed as a function of the flavor fractions (the unknown quantities) and the $b-$tagging probabilities. To reduce the number of unknown flavor fractions from 6 to 3,

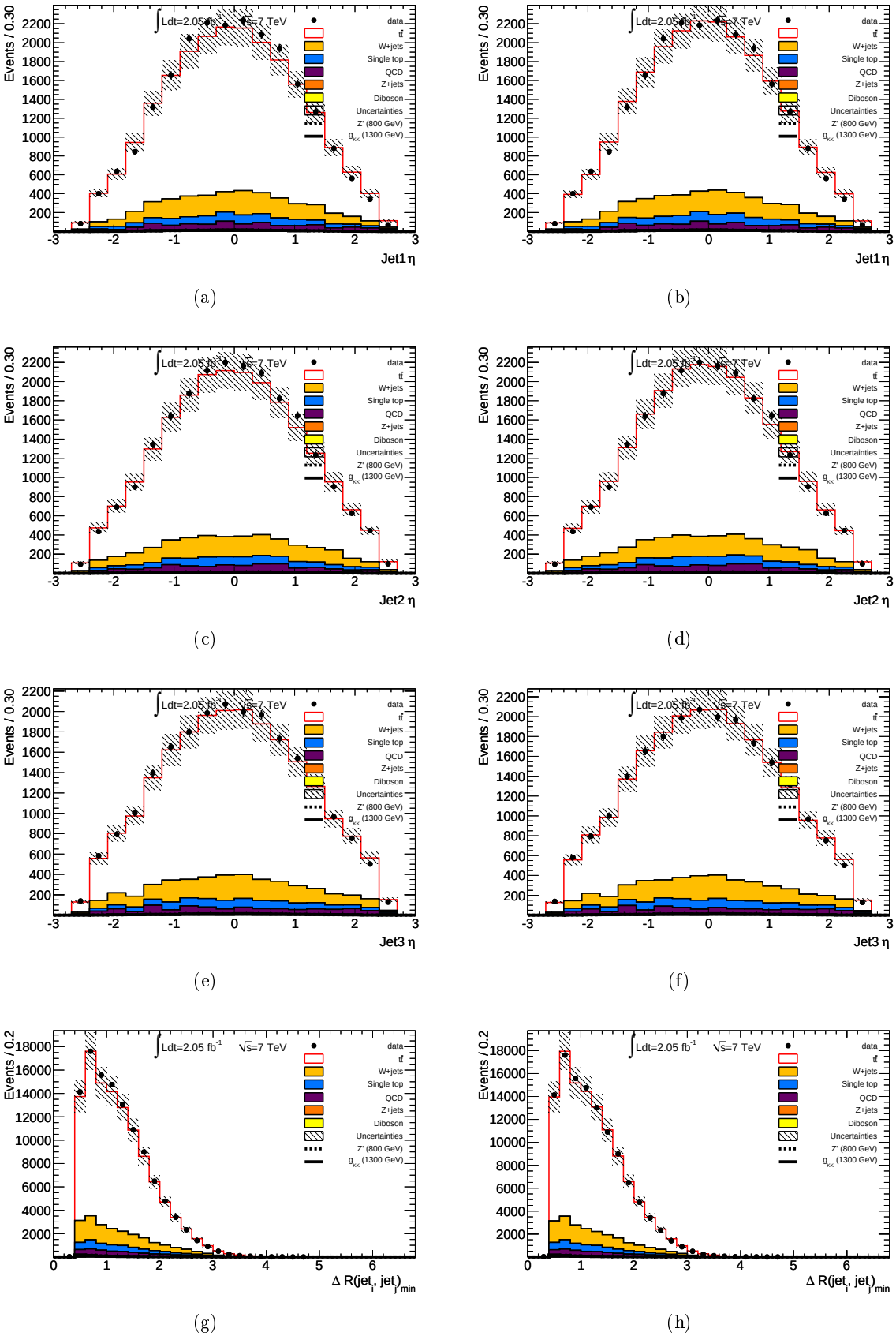
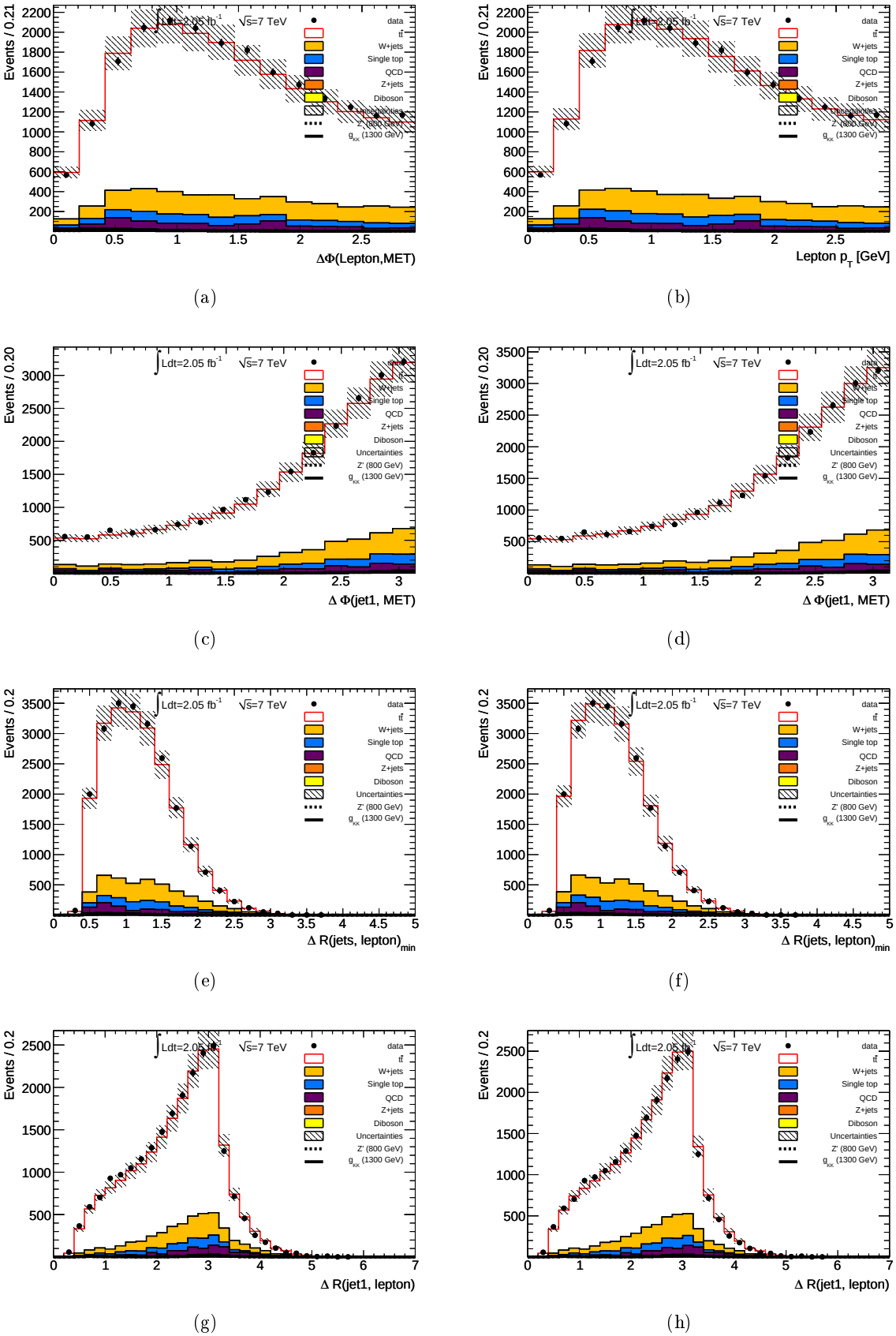


Figure 6.5: (a,b) Leading p_{T} jet η, (c,d) second jet η, (e,f) third jet η and (g,h) minimum ΔR between jets in the event, $\Delta R_{min}(jet, jet)$ when using CTEQ6.6 (left) and MRST2008NLO (right) for the SM $t\bar{t}$ and single top samples. e and μ plus jets channels are combined. The grey hashed are shows the total background normalization uncertainty (which will be fully described in Section 6.10.1).



(a) (b)

(c) (d)

(e) (f)

(g) (h)

Figure 6.6: (a,b) $\Delta\phi$ between the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$, (c,d) $\Delta\phi$ between the leading jet and the $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(jet1, E_{\mathrm{T}}^{\mathrm{miss}})$, (e,f) minimum ΔR between jets in the event and the lepton, $\Delta R_{min}(\ell, jets)$, and (g,h) ΔR between the leading jet and the lepton, $\Delta R(\ell, jet1)$, when using CTEQ6.6 (left) and MRST2008NLO (right) for the SM $t\bar{t}$ and single top samples. e and μ plus jets channels are combined. The grey hashed are shows the total background normalization uncertainty (which will be fully described in Section 6.10.1).

the ratio of the flavor fractions between the 1-jet and the 2-jets bin calculated in simulation is used. In addition the ratio between the $W + b\bar{b}$+jets and $W + c\bar{c}$+jets in the 2-jets bin is fixed to the value estimated from simulation too, to reduce the system to two equations with two unknown variables. Once the flavor fraction have been calculated in the 2-jets bin, they can be extrapolated to a higher multiplicity bin using the ratio of heavy flavour fractions in that given bin estimated from Monte Carlo. Based on these studies the $W+b\bar{b}$ and $W+c\bar{c}$ components were scaled by a factor of 1.63, the $W + c$ by 1.11 and the light component by 0.83. The scaling for the light component was calculated in such a way that the overall normalization of the W+jets samples remained unchanged in the 2-jets bin.

The flavour composition uncertainty of the W+jets background is estimated by varying these scaling factors by their statistical uncertainties (13% for $W + b\bar{b}$ and $W + c\bar{c}$, 9% for $W + c$). An additional 25% is added in quadrature to account for the extrapolation to a higher jet multiplicity. Uncertainties on jet energy scale, b-tagging efficiency, etc. are applied in addition to these W+jets-specific uncertainties. The latter uncertainties are of course correlated with the uncertainties in the analysis. To properly include this correlations, the flavor composition W+jets scale factors were recalculated for each uncertainty scenario.

6.7.3 Multijet background

Fakes leptons in the multijet background can originate from different processes, for example non-prompt leptons from heavy flavor quark decays, π^0s misidentified as electrons or photon conversions. These fake leptons may pass the analysis lepton selection, eventhough they were not produced in a W boson decay. The multijet background in a $b-$tagged analysis as this one is mainly dominated by non-prompt leptons in both channels. The generation of the multijet background in simulation is a difficult task. Multijet processes have a large cross section. We are only interested in the small multijet phase-space where fake leptons are produced. Many multijet events would need to be generated in order to get a statistically meaningful simulated multijet sample for the analysis. To avoid these inconvenients the multijet background kinematic distributions and its normalization can be obtained from data. The multijet background is modeled using a data-derived template, which is fitted to the selected data in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution to estimate the multijet background expectation. Due to the similarities in the multijet background composition in the electron and muon channels the same data-derived templates are used.

The method used to estimate the multijet background is known as the jet-electron method [157]. It consists in selecting events with kinematic characteristics similar to the events that pass all the analysis lepton selection criteria, but using a jet with a high EM fraction instead of an electron. A data-derived sample is collected using events triggered by a jet trigger containing exactly one jet with a high EM fraction (between 0.8 and 0.95). This jet is also required to have at least 4 tracks, to avoid contamination from photon conversions. A veto on good electrons of medium quality is applied. This sample is used to model the kinematic distribution shapes for the multijet background. It is obtained before $b-$tagging to reduce statistical fluctuations.

On the other hand, to calculate the corresponding normalization a binned likelihood fit of the jet-electron sample together with the SM $t\bar{t}$, single top, W and Z+jets simulated samples is performed to the data using the full $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. The fit is done after applying all the selection criteria except the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. The diboson is not included, since its contribution is negligible in the analysis. The multijet background and the signal $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution are different enough so that fitting the multijet contribution to the full $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution will not mask a potential signal. The simulated samples for SM $t\bar{t}$, single top, W and Z+jets contributions are allowed to vary in each $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution bin according to a gaussian distribution centered at the bin height, with 10% RMS to account for their own modeling uncertainties. Figure 6.7 and

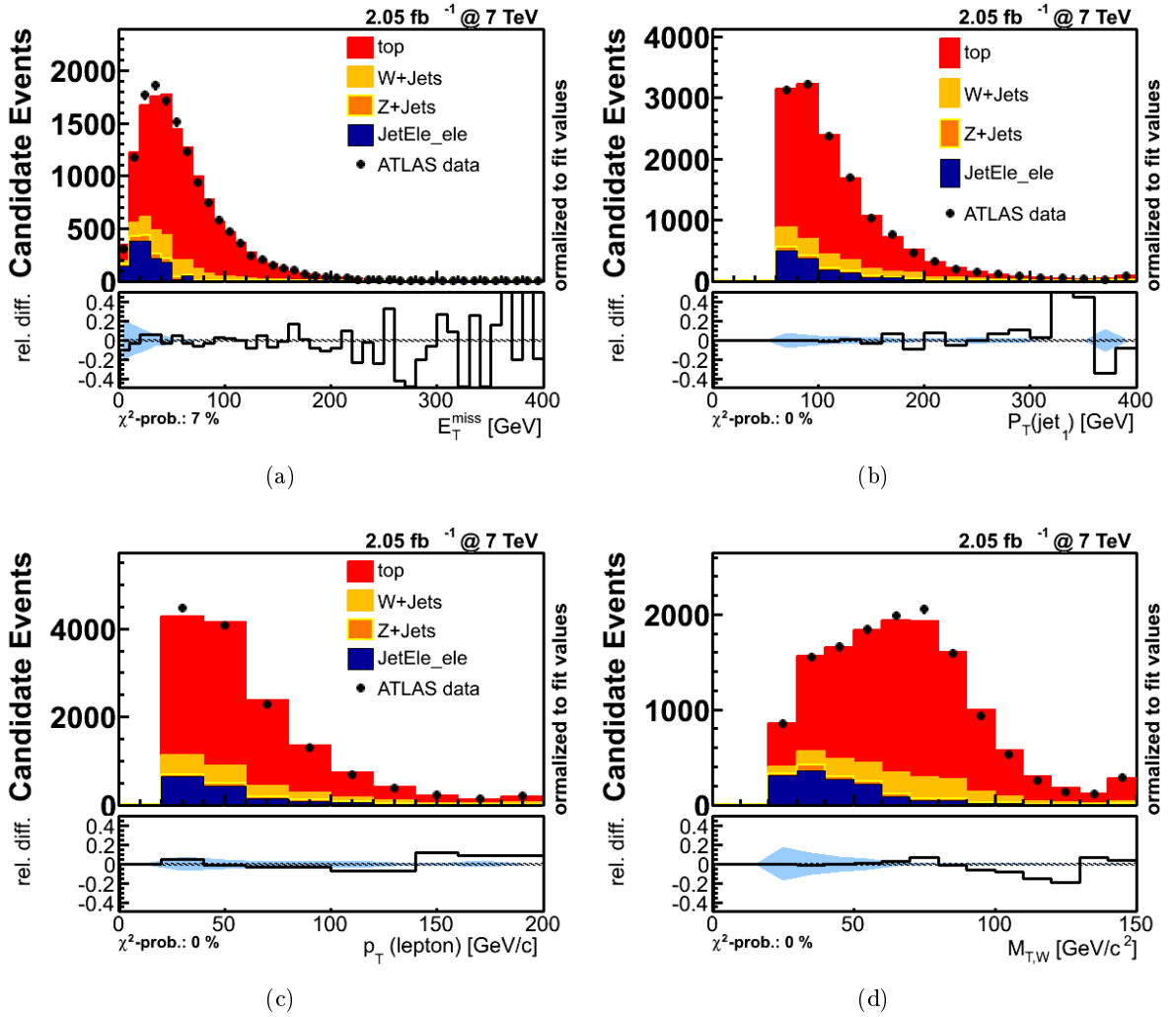


Figure 6.7: (a) $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution used for the estimation of the multijet background in the electron channel after $b-$tagging, using the jet-electron method; (b) leading jet p_{T}, (c) lepton p_{T} and (d) lepton-$E_{\mathrm{T}}^{\mathrm{miss}}$ transverse mass distributions after b-tagging. The shaded blue areas in the relative difference plots show the 50% systematic uncertainty assigned to the multijet background normalization. No cut in the $E_{\mathrm{T}}^{\mathrm{miss}}$ is applied [8].

6.8 show the $E_{\mathrm{T}}^{\mathrm{miss}}$, leading lepton p_{T}, the leading jet p_{T} and the W transverse mass distributions after the fit in the electron and the muon channels, respectively. The $E_{\mathrm{T}}^{\mathrm{miss}}$ cut was not applied.

Two kinds of systematic uncertainties are applied. The first one is a normalization uncertainty. It is calculated by comparing the standard fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution described above with similar fits using the W transverse mass distribution and fits to the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for different ranges of primary vertices multiplicity. Maximum fit differences of 17.0% and 48.3% for the electron and muon channel, respectively, are found. Therefore, the normalization systematic uncertainty has been set to 50%. The second ones are two shape systematic uncertainties. They are estimated by comparing the shapes of the reconstructed $t\bar{t}$ mass spectra for different pile up levels ($<$ and ≥ 6 reconstructed primary vertices) and for different ranges in the transverse energy of the event H_{T}. H_{T} corresponds to the scalar sum of transverse momentum of the jets, the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event. To give equal statistical weights to both subsamples the separation is set at $H_{\mathrm{T}} = 350$ GeV and 310 GeV in the electron and muon channels, respectively.

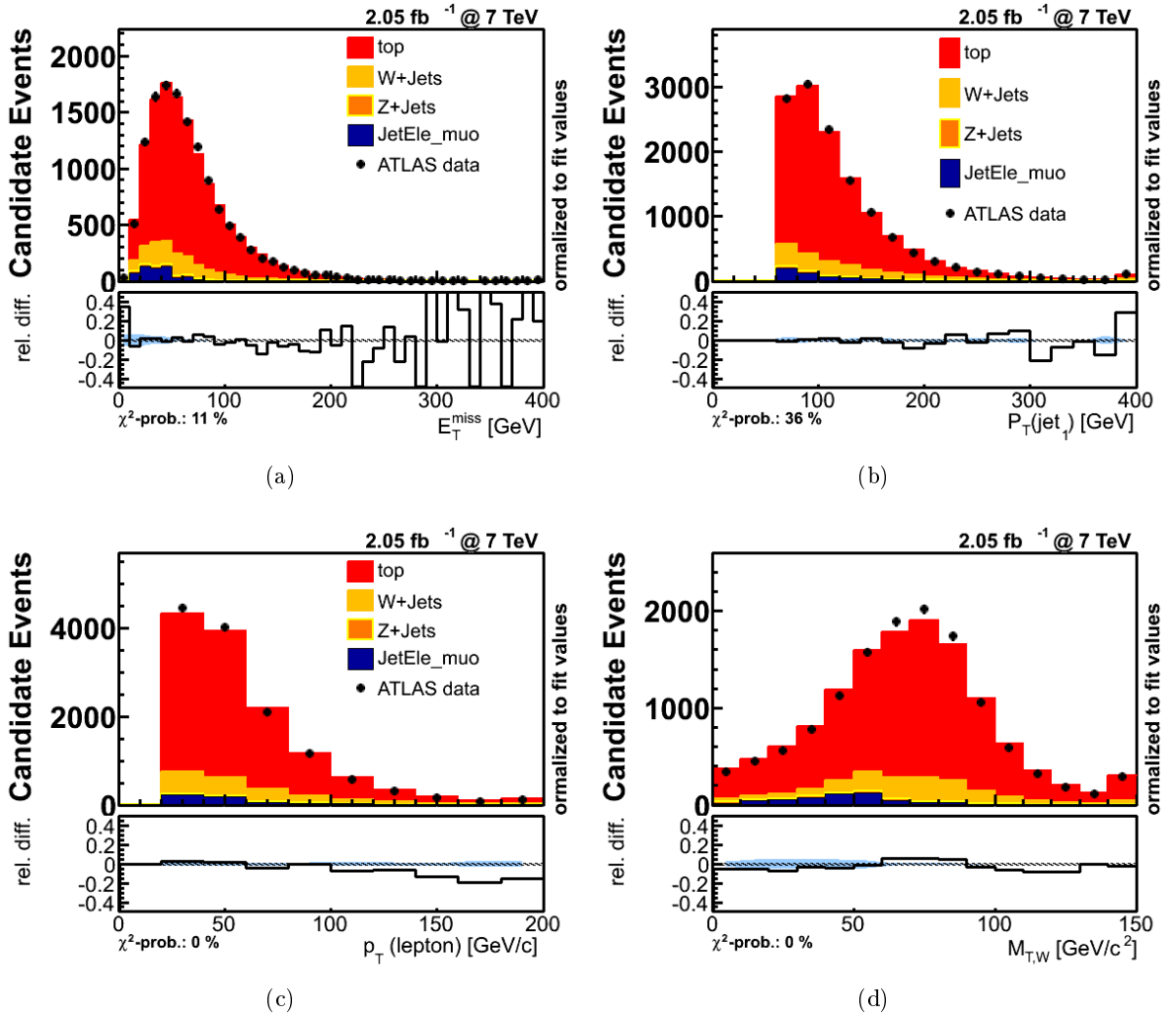


Figure 6.8: (a) $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution used for the estimation of the multijet background in the muon channel after $b-$tagging, using the jet-electron method; (b) leading jet p_{T}, (c) lepton p_{T} and (d) lepton-$E_{\mathrm{T}}^{\mathrm{miss}}$ transverse mass distributions after b-tagging. The shaded blue areas in the relative difference plots show the 50% systematic uncertainty assigned to the multijet background normalization. No cut in the $E_{\mathrm{T}}^{\mathrm{miss}}$ is applied [8].

6.8 Data versus background expectation comparison

Table 6.8 shows the event yields for data and background expectation after applying the event selection described in Section 6.4 and the PDF reweighting of the $t\bar{t}$ and single top samples. Event and reconstructed objects corrections discussed in Chapters 3 and 4 are applied. The multijet and the W+jets background are estimated as described in Section 6.7. Figures 6.9 and 6.10 show the distribution for some kinematic variables in the event for data, the different backgrounds considered in the analysis and two signal points, a Z' with $m_{Z'} = 800$ GeV and a KK gluon with $m_{g_{KK}} = 1300$ GeV. Both channels have been combined in these plots. These kinematic distributions are shown separately for the muon and the electron channel in Annexes E and D, respectively. The grey hashed area correspond to the total background normalization systematic uncertainty (which will be fully described in Section 6.10.1). A good agreement between data and background expectation has been found within the uncertainties. In Figures 6.9c and 6.9f the

	Electron channel	Muon channel
$t\bar{t}$	7830.9	10006.3
Single top	466.4	569.4
W+jets	1119.1	1448.9
Z+jets	83.5	87.7
Diboson	17.6	17.6
Total MC	9517.5	12129.9
Multijet	344.0	471.4
Total Expected	9861.6	12601.2
Data observed	9622	12706
Z', m = 1000 GeV	24.5	26.0
g_{KK}, m = 1300 GeV	58.9	65.1

Table 6.8: Number of expected signal and background and observed events for the electron and muon channels after applying all selection cuts [8].

last bin includes the overflow above 1 TeV. In Figures 6.10a and 6.10b the number of $b-$tagged jets for the multijet background is zero since the jet-electron template was generated before $b-$tagging as explained in Section 6.7.3. This also explains the fact that the `JetFitterCombNN` weight is set to -1 for the multijet background as shown in Figure 6.10c.

6.9 $t\bar{t}$ pair reconstruction

The discriminant variable used in the search for $t\bar{t}$ resonances in the lepton plus jets channel is the invariant mass of the $t\bar{t}$ pair. If a new heavy resonance exists it should appears as a bump in the invariant mass spectra of the $t\bar{t}$ pair. Several methods have been used to reconstruct the invariant mass of the $t\bar{t}$ pair. Two different approaches can be followed, either reconstructing the individual top quarks or not. The χ^2 algorithm falls in the first category. While the four hardest jets and dRmin methods fall in the second one. The $t\bar{t}$ pair is built by summing up the four momentum of the selected jets, the lepton and the neutrino. Before discussing the different $t\bar{t}$ reconstruction methods, the neutrino's reconstruction will be described in Section 6.9.1. All methods have been adapted in case of a possible partially resolved topology, i.e. high mass events.

6.9.1 Neutrino's reconstruction

In $t\bar{t}$ lepton plus jets events, only one neutrino is supposed to be produced from one of the W bosons decay. $E_{\mathrm{T}}^{\mathrm{miss}}$ only defines the x and y components of the neutrino's momentum, but no information about the z component is provided. The W boson mass, $M_W = 80.4$ GeV [13], can be used as a constraint to obtain an equation for the z component of the neutrino, $p_{z,\nu}$. The sum of the four-vectors of the lepton, p_l, and the neutrino, p_ν, is equal to the four-vector of the W boson, p_W:

$$\mathbf{p}_W = \mathbf{p}_l + \mathbf{p}_\nu. \tag{6.2}$$

After some manipulations, Eq. 6.2 can be written as:

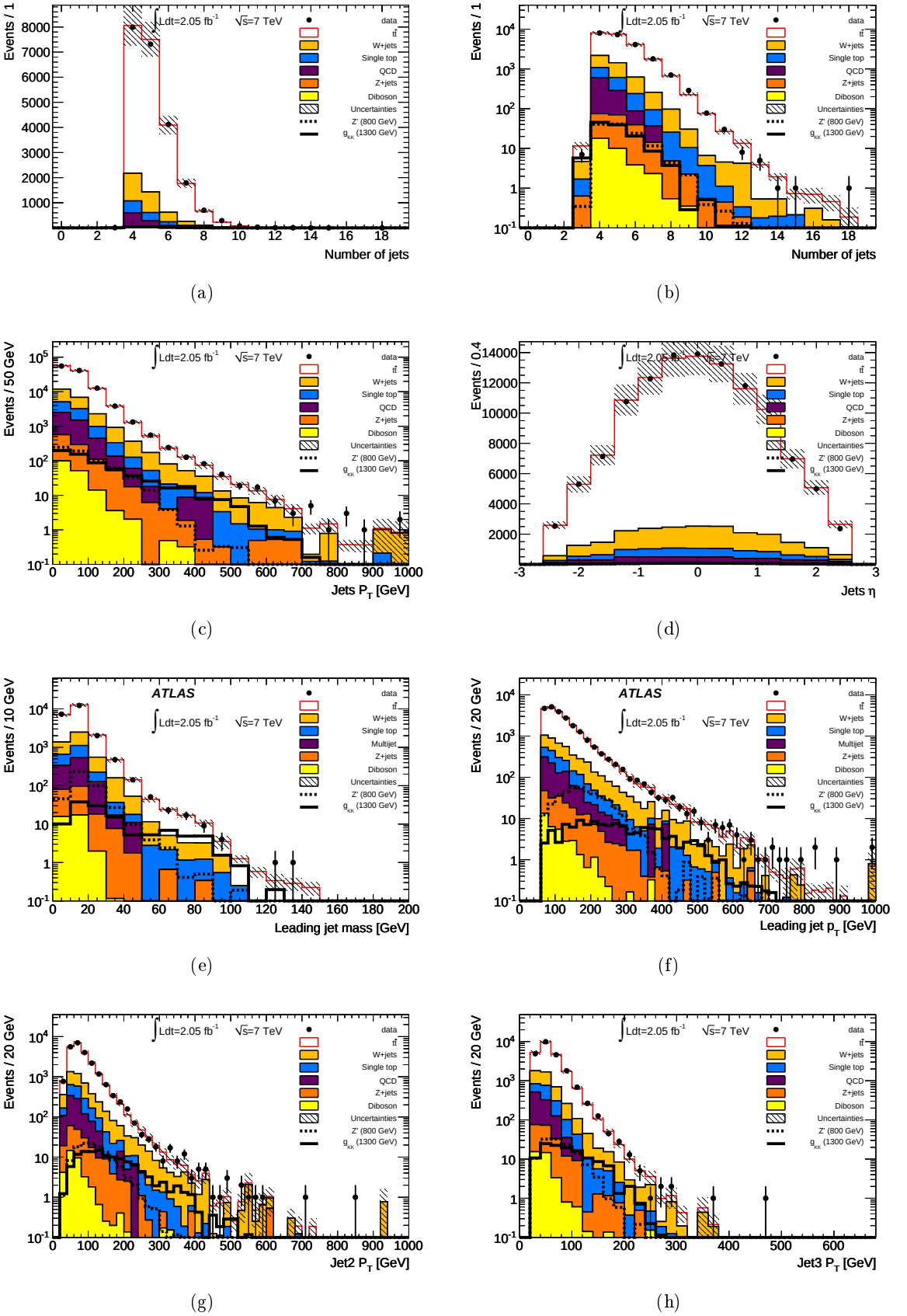


Figure 6.9: Number of jets with $p_{\mathrm{T}} > 20$ GeV in (a) linear and (b) logarithmic scales, jets (c) $p_{\mathrm{T}}$and (d) η distributions, (e) leading jet mass, (f) leading jet p_{T}, (g) second jet p_{T} and (h) third jet p_{T} after all selection criteria were applied. The grey hashed area shows the total background normalization uncertainty.

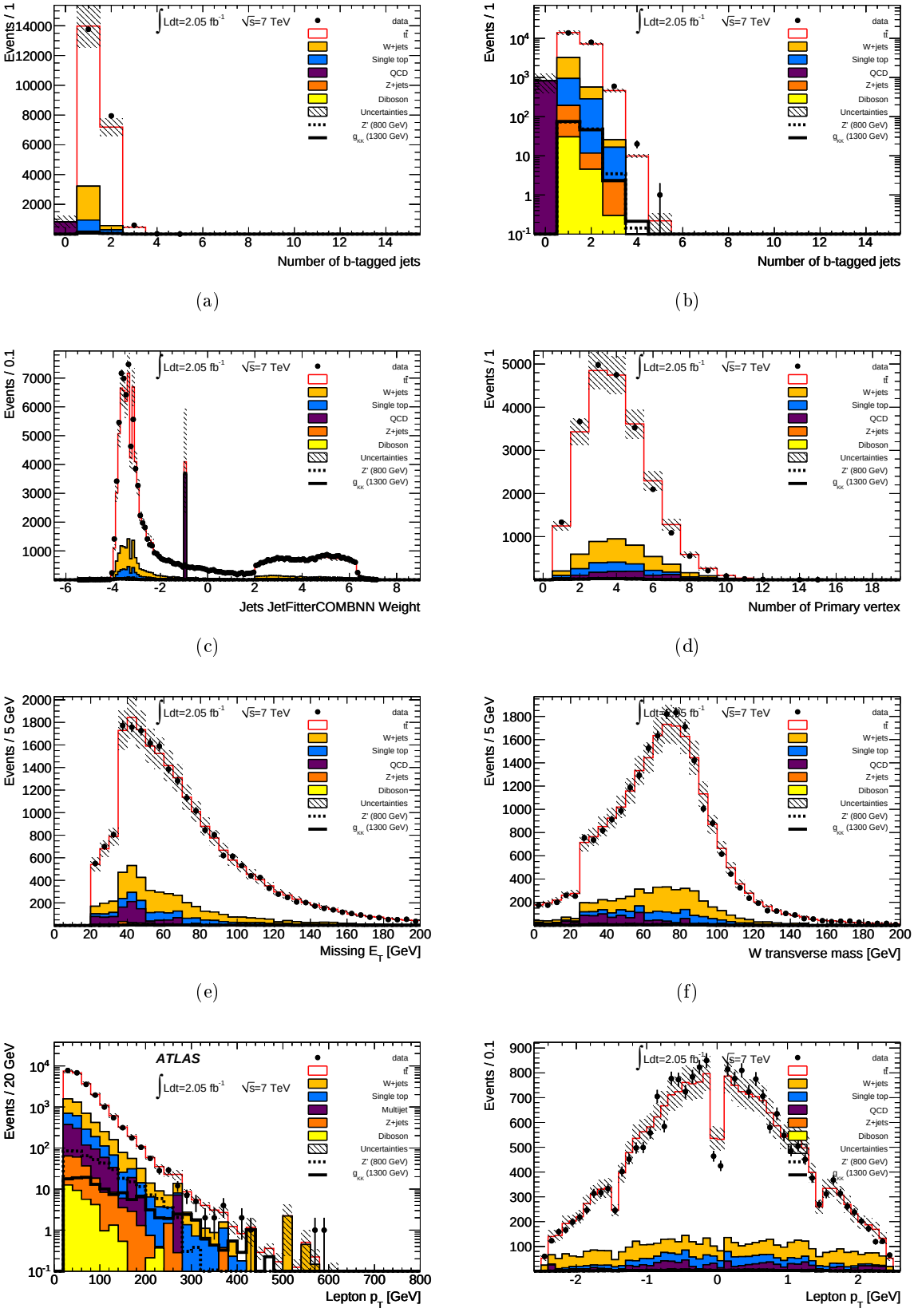


Figure 6.10: Number of $b-$tagged jets with $p_{\mathrm{T}} > 20$ GeV in (a) linear and (b) logarithmic scales, (c) JetFitterCOMBNN weight for all jets, (d) number of primary vertices, lepton (e) p_{T} and (f) η, (g) $E_{\mathrm{T}}^{\mathrm{miss}}$ and (h) W boson transverse mass after all selection criteria were applied. The grey hashed area shows the total background normalization uncertainty.

$$p_{z,\nu}^2 - 2\frac{\mu p_{z,l}}{E_l^2 - p_{z,l}^2} p_{z,\nu} + \frac{E_l^2 p_{\mathrm{T},\nu}^2 - \mu^2}{E_l^2 - p_{z,l}^2} = 0 \tag{6.3}$$

$$\mu = \frac{1}{2} M_W^2 + p_{\mathrm{T},l} p_{\mathrm{T},\nu} \cos\Delta\phi, \tag{6.4}$$

where $\Delta\phi$ is the azimuthal angle between the charged lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$, E_l is the charged lepton energy, $p_{z,l}$ is the z component of the charged lepton momentum and $p_{\mathrm{T},\nu}$ is the transverse component of the neutrino momentum, which corresponds to the $E_{\mathrm{T}}^{\mathrm{miss}}$. Eq. 6.4 can be solved for $p_{z,l}$:

$$p_{z,\nu}^{\pm} = \frac{\mu p_{z,l}}{p_{\mathrm{T},l}} \pm \sqrt{\frac{\mu^2 p_{z,l}^2}{p_{\mathrm{T},l}^4} - \frac{E_l^2 p_{\mathrm{T},\nu}^2 - \mu^2}{p_{\mathrm{T},l}^2}}. \tag{6.5}$$

If the discriminant in Eq. 6.5 is positive then two real solution are obtained for $p_{z,\nu}$. On the other hand, if it is negative, two complex solutions are obtained. The discriminant becomes negative when the reconstructed transverse mass of the W boson, $m_{\mathrm{T,W}}$, is bigger than M_W. The main reason of this behavior has been shown to be the imperfect resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ [158]. In the first case when the discriminant is positive, the smallest $|p_{\mathrm{z}}|$ solution has been shown to give the best neutrino p_{z} resolution, as can be seen in Figure 6.11a. In the second case when the discriminant is negative, one possibility is to take only the real part of Eq. 6.5 to calculate $p_{z,\nu}$. Another possibility is to adjust the $E_{x,y}^{\mathrm{miss}}$ components to get a null discriminant, by allowing small rotations of the $E_{\mathrm{T}}^{\mathrm{miss}}$ to satisfy $m_{\mathrm{T,W}} = M_W$ [158]. Only one solution is obtained in this case. The neutrino momentum resolution for events with a negative discriminant, when each one of the above solutions are used, is shown in Figure 6.11.

6.9.2 Four Hardest Jets

The four hardest jets is the simplest method used to reconstruct the $t\bar{t}$ pair. In the case of two solutions for the neutrino's p_{z}, the neutrino is reconstructed using the smallest $|p_{\mathrm{z}}|$, which has shown to give the best neutrino p_{z} resolution (see Figure 6.11a). In the case of a resolved topology, the $t\bar{t}$ pair is reconstructed by combining the four highest p_{T} jets, the selected charged lepton and the reconstructed neutrino. For high mass events using the four highest p_{T} jets with $|\eta| < 2.5$ could lead to an overestimation of $m_{t\bar{t}}$. Therefore, only the three highest p_{T} jets are used in the mass reconstruction if one of the jets has mass $m^{\mathrm{jet}} > 60$ GeV. Figures 6.12a and 6.12b show the reconstructed invariant masses and corresponding resolutions obtained using the 4 hardest jets algorithm for four different resonance masses: $m_{Z'} = 500$, 700, 1000 GeV and $m_{g_{KK}} = 1300$ GeV. The reconstructed $t\bar{t}$ pair invariant mass is shown as a function of the true mass for SM $t\bar{t}$ production and Z' with $m_{Z'} = 2000$ GeV in Figures 6.13a and 6.14a, respectively.

The dominant source of long, non-gaussian tails in the mass resolution is the use of a jet from ISR or FSR in the place of one of the jets comming directly from the top-quark decay. Using an ISR jet will lead to an overestimated reconstructed $t\bar{t}$ invariant mass because of the large angle w.r.t. the initial parton, while using a FSR jet will lead to an underestimated reconstructed $t\bar{t}$ invariant mass. There are three more effects that can play a role in the $t\bar{t}$ pair reconstruction. The first one is also related with the FSR. The gluons emitted will take away part of the energy of the initial partons. If they are not taken into account to reconstruct the $t\bar{t}$ pair, the resulting reconstructed invariant mass is underestimated. This radiation emission increases when the energy of the initial quarks increases. Therefore this effect becomes more important for higher resonance masses. The second effect is related with the event pile-up. With increasing pile-up,

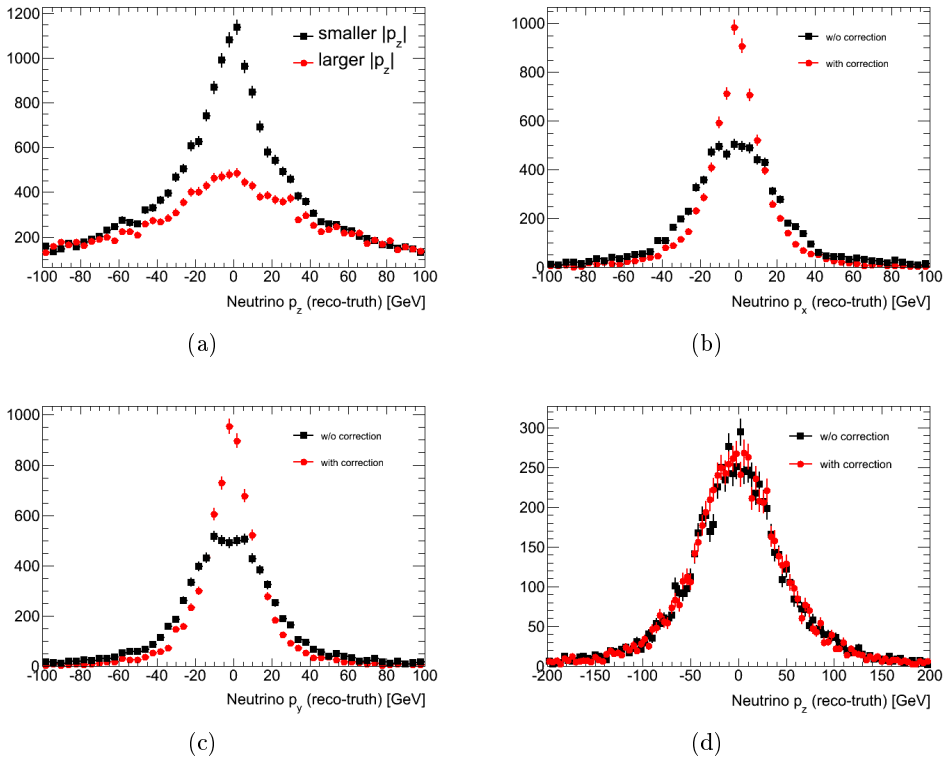


Figure 6.11: (a) Comparison of the neutrino p_z resolution for the smaller and larger solution in the case of two solutions to the quadratic equation, (b) $p_{x,\nu}$, (c) $p_{y,\nu}$ and (d) $p_{z,\nu}$ resolution for events with a negative discriminant in Eq. 6.5, where only the real part of the solution is taken (black) and the $E_{\mathrm{T}}^{\mathrm{miss}}$ is adjusted to get a null discriminant (red) [8].

the number of reconstructed jets increases. Using a jet from pile-up in the place of one of the jets coming from the tops decays will lead usually to a underestimated reconstructed $t\bar{t}$ pair invariant mass. The third one regards only resonances with a high mass with a merged topology. If the $t\bar{t}$ pair is reconstructed with four jets instead of three or even two jets, then $m_{t\bar{t}}$ will be overestimated.

The reconstruction efficiency as a function of the Z' and g_{KK} mass is shown in Figure 6.15. The reconstruction efficiency has been defined only for events without a high mass jet, as the fraction of matchable events where four jets selected by the reconstruction algorithm correspond to the reconstructed jets matched to the partons coming from the top-quarks decay. A matchable event is a $t\bar{t}$ event with a lepton plus jets topology for which each parton coming from the tops is matched to an unique reconstructed object. A given reconstructed object can not be matched to several partons. The reconstruction efficiency for events with a high mass jet will be included in the next iteration.

6.9.3 dRmin

The dRmin method is a variant of the four hardest jet method. Its main goal is to reduce the probability of selecting jets from ISR in the $t\bar{t}$ reconstruction instead of jets coming from top-quark decay, by requiring a topological proximity between the selected jets. The dRmin algorithm considers the four highest p_{T} jets with $|\eta| < 2.5$, and a jet is excluded if its angular distance to the lepton or closest jet satisfies $\Delta R > 2.5 - 0.015 \times m^{\mathrm{jet}}$, where m^{jet} is the jet's mass. If more than one jet satisfies this condition, the jet with the largest ΔR is excluded. If a

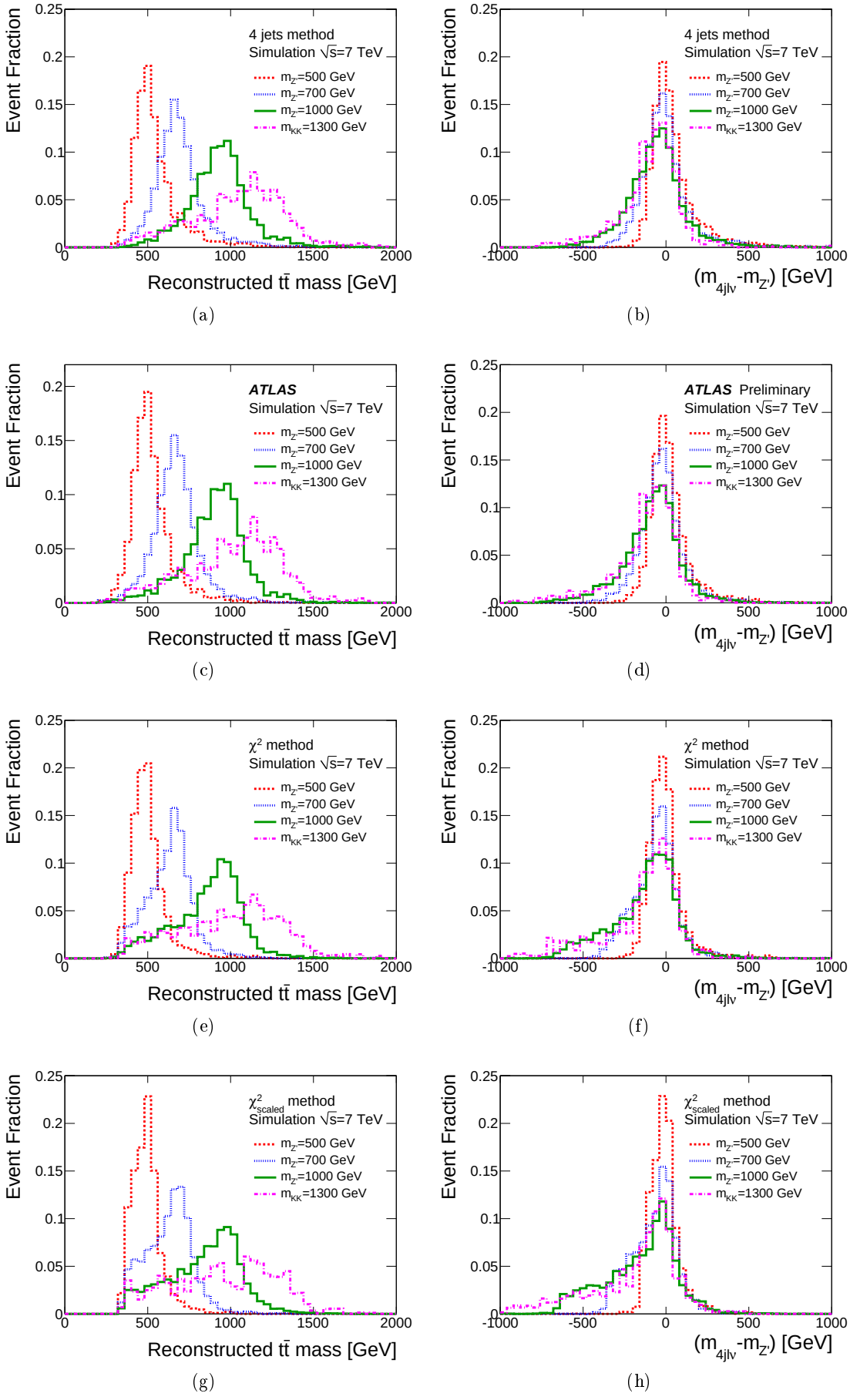


Figure 6.12: Reconstructed $t\bar{t}$ pair invariant mass (left) and resolution on the reconstructed $t\bar{t}$ pair invariant mass (right) using the (a,b) four hardest jet, (c,d) dRmin, (e,f) χ^2 and (g,h) scaled χ^2 method for four different resonance masses: $m_{Z'} = 500, 700, 1000$ GeV and $m_{g_{KK}} = 1300$ GeV.

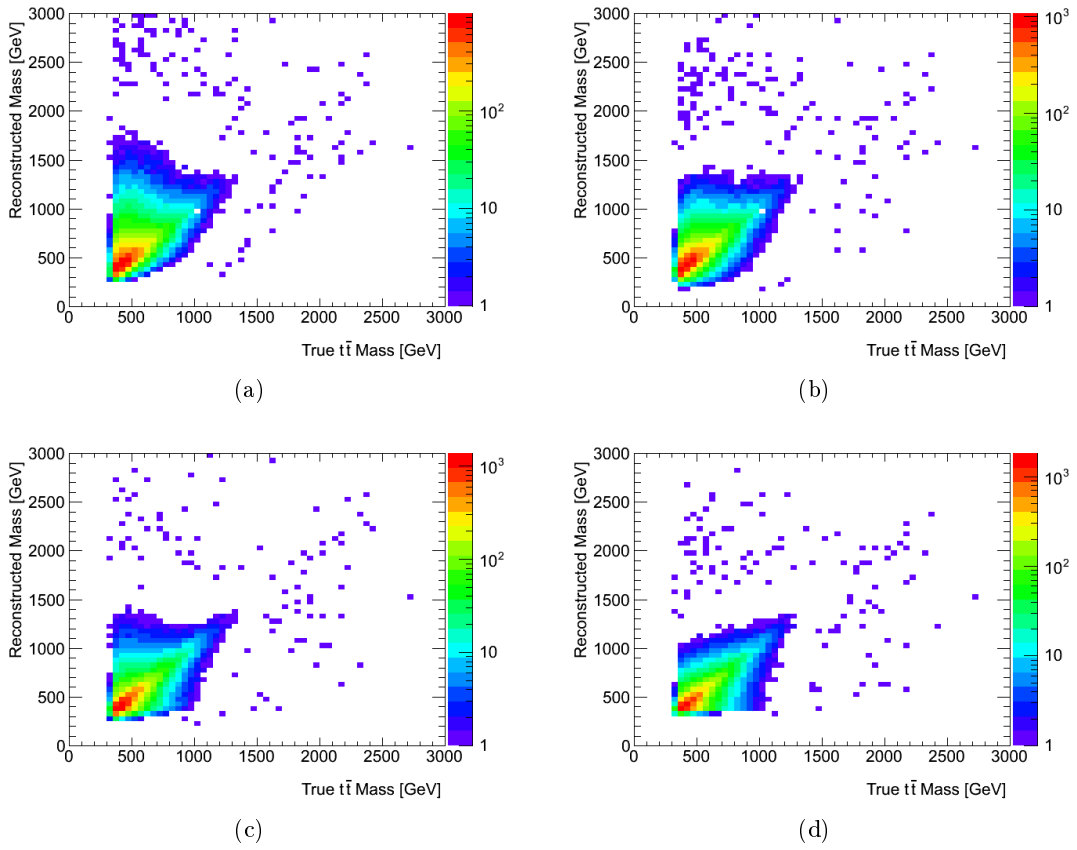


Figure 6.13: Reconstructed versus true $t\bar{t}$ pair invariant mass using the (a) four hardest jets, (b) dRmin, (c) χ^2 and (d) scaled χ^2 method for SM $t\bar{t}$ production.

jet is discarded and more than 3 jets remain, the procedure is repeated. Figure 6.16 shows the correlation between the ΔR to the closest jet and m^{jet} for both jets matched and not matched to top-quark decay products in SM $t\bar{t}$ events. A good agreement between data and simulation in the angular distance to the lepton or closest jet distributions is observed in Figures 6.5 and 6.6. The $t\bar{t}$ pair is reconstructed from the lepton, the neutrino and the leading four jets, or three jets if only three remain. To reconstruct the neutrino, the smallest $|p_{\text{z}}|$ solution (in the case of a positive discriminant) is used (see Figure 6.11a).

By allowing only three jets in the final state the dRmin method deals with the possibility that one of the jets from the top-quark decay to be outside the detector acceptance or merged with another jet. In the case of high mass events, potentially merged decay products are handled in a slightly different way. If one of the jets has mass $m^{\text{jet}} > 60$ GeV, it is combined with the closest jet in ΔR to form the hadronic top quark candidate. The leptonic top quark is reconstructed by combining the reconstructed leptonic W boson candidate with the closest jet in ΔR. The reconstructed invariant masses and corresponding resolutions obtained using the dRmin method is presented in Figures 6.12c and 6.12d for four different resonance masses: $m_{Z'} =$ 500, 700, 1000 GeV and $m_{g_{KK}} = 1300$ GeV. The reconstructed $t\bar{t}$ pair invariant mass is shown as a function of the true mass for SM $t\bar{t}$ production and Z' with $m_{Z'} = 2000$ GeV in Figures 6.13b and 6.14b, respectively. In Figure 6.12 a reduction of the right tails is observed for the resolution plots with respect to the results obtained using the four hardest jet algorithm. In addition, the reconstruction efficiency is slightly improved as can be see in Figure 6.15.

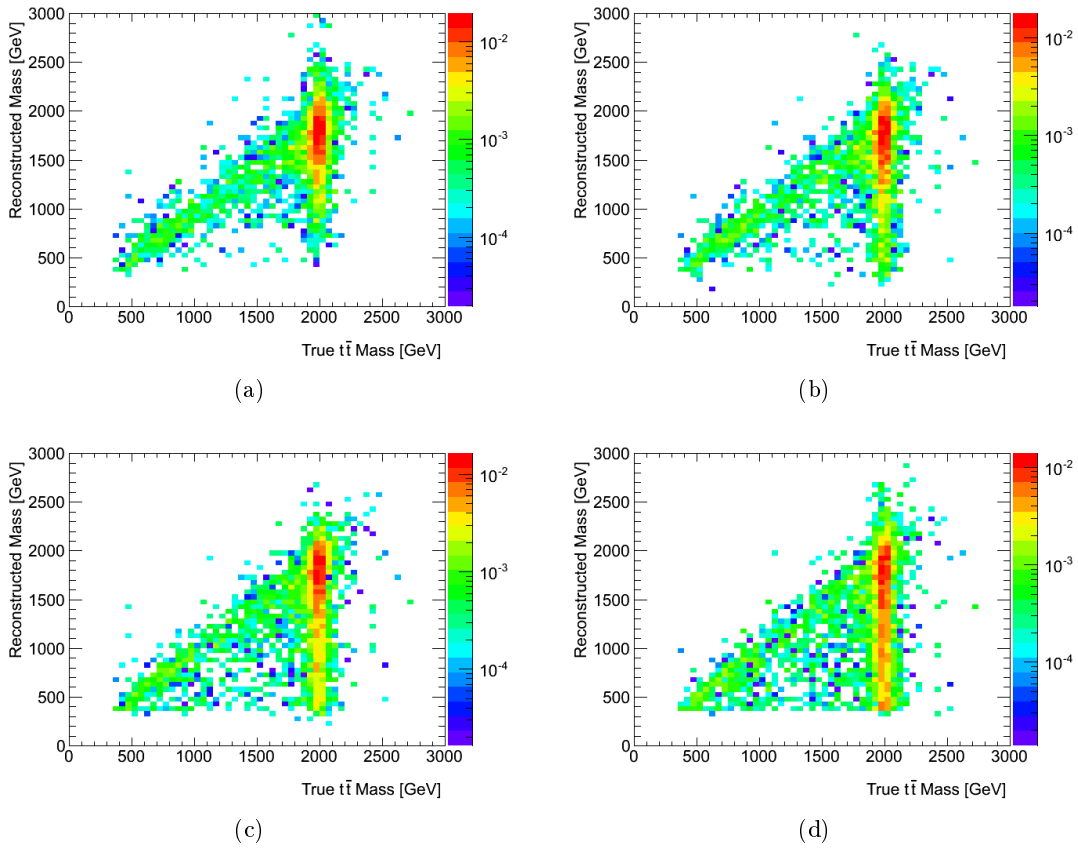


Figure 6.14: Reconstructed versus true $t\bar{t}$ pair invariant mass using the (a) four hardest jets, (b) dRmin, (c) χ^2 and (d) scaled χ^2 method for Z' with $m_{Z'} = 2000$ GeV.

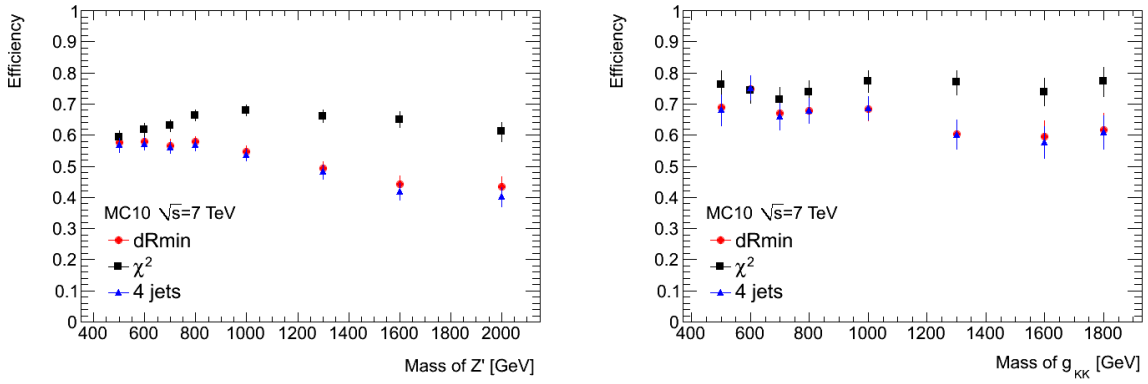
6.9.4 χ^2 algorithm

The χ^2 algorithm reconstructs the individual top quarks to calculate $m_{t\bar{t}}$. To reduce the tails in the invariant mass reconstruction from wrong assignments, a χ^2 is constructed using the known top-quark and W boson mass as constraints:

$$\chi^2 = \left[\frac{m_{jj} - m_W}{\sigma_W}\right]^2 + \left[\frac{m_{jjb} - m_{jj} - m_{th-W}}{\sigma_{th-W}}\right]^2 + \left[\frac{m_{j\ell\nu} - m_{tl}}{\sigma_{t\ell}}\right]^2 + \left[\frac{(p_{\mathrm{T},jjb} - p_{\mathrm{T},j\ell\nu}) - (p_{\mathrm{T},th} - p_{\mathrm{T},t\ell})}{\sigma_{p_{\mathrm{T},th} - p_{\mathrm{T},t\ell}}}\right]^2. \tag{6.6}$$

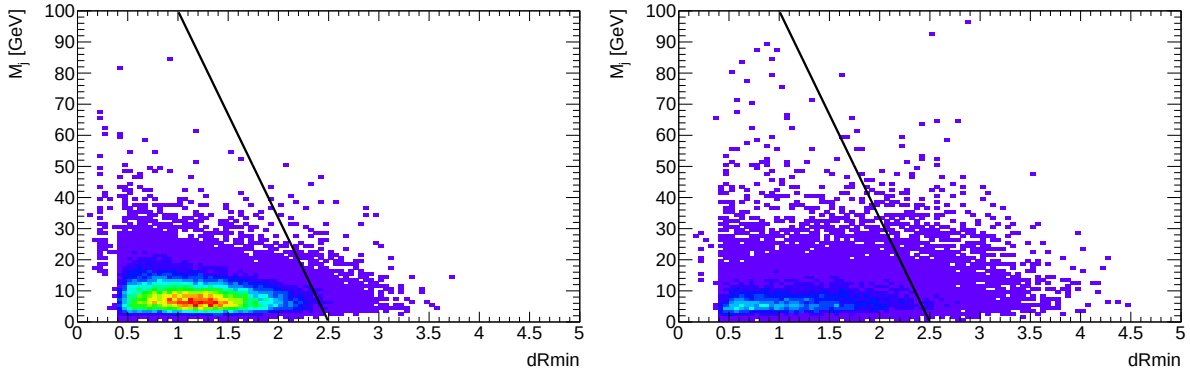
The first term is the constraint from the hadronically decaying W boson. The second term corresponds to the hadronically decaying top quark, but since m_{jj} and m_{jjb} are heavily correlated the hadronically decaying W-boson was subtracted, $th - W$, to decouple this term from the previous one. The third term represents the top-quark leptonic decay, tl. The last term constrains the difference between the two top quarks transverse momentum, $p_{\mathrm{T},th} - p_{\mathrm{T},t\ell}$. For a resonance decay the difference of transverse momentum between the hadronic and leptonic top is expected to be close to zero, if the resonance is produced with low p_{T}. Figure 6.17 shows this difference calculated using matched events.

m_W, m_{th-W}, $m_{t\ell}$, $p_{\mathrm{T},th} - p_{\mathrm{T},t\ell}$ and their respective resolutions were determined using the reconstructed objects matched to the partons coming from the top quarks in matchable events. To derive the parameters values Z' samples were used for masses between $m_{Z'} = 700$ GeV to 2000 GeV. It allows to reduce the statistical errors and to gain sensitivity to high mass resonances. Their values were found to be: $m_W = 84.0$ GeV, $m_{th-W} = 89.9$ GeV, $m_{t\ell} = 166.7$ GeV, $\sigma_W =$



(a) (b)

Figure 6.15: Reconstruction efficiency using the four hardest jets, dRmin, χ^2 (same for χ^2 scaled) methods as a function of the (a) Z' and (b) g_{KK} mass.



(a) (b)

Figure 6.16: Correlation between the angular separation to the closest jet and jet mass for jets a) matched and b) not matched to top-quark decay products. Jets to the right of the black line are rejected. The absolute color scale is the same in both plots [8].

11.1 GeV, $\sigma_{th-W} = 12.7$ GeV, $\sigma_{t\ell} = 23.9$ GeV, $p_{\mathrm{T},th} - p_{\mathrm{T},t\ell} = -3.4$ GeV and $\sigma_{diff p\mathrm{T}} = 36.9$ GeV. All possible jet permutations and neutrino's p_z solutions were tried and only the permutation with lowest χ^2 was used. At least one $b-$tagged jet was required in the combination.

If one of the jets has mass $m^{\mathrm{jet}} > 60$ GeV, the χ^2 is changed to be:

$$\chi^2 = \left[\frac{m_{jj} - m_{jj}^{th}}{\sigma_{jj}^{th}}\right]^2 + \left[\frac{m_{j\ell\nu} - m_{t\ell}}{\sigma_{t\ell}}\right]^2 + \left[\frac{(p_{\mathrm{T},jj} - p_{\mathrm{T},j\ell\nu}) - (p_{\mathrm{T},th} - p_{\mathrm{T},t\ell})}{\sigma_{p_{\mathrm{T},th} - p_{\mathrm{T},t\ell}}}\right]^2, \tag{6.7}$$

where the $m_{jj} - m_{jj}^{th}$ term allows the merging of either both quarks from W boson decay, or one quark from W boson decay with the b quark from top-quark decay. Similarly, the values of m_{jj}^{th} and σ_{jj}^{th} were determined from simulation to be 167.2 GeV and 18.6 GeV, respectively. At least one $b-$tagged jet was required in the combination. Figures 6.12e and 6.12f show the reconstructed mass and the corresponding mass resolution for four different resonance masses: $m_{Z'} = 500$, 700, 1000 and $m_{g_{KK}} = 1300$ GeV. The reconstructed $t\bar{t}$ mass is shown as function of the true mass for SM $t\bar{t}$ production and Z' with $m_{Z'} = 2000$ GeV in Figures 6.13c and

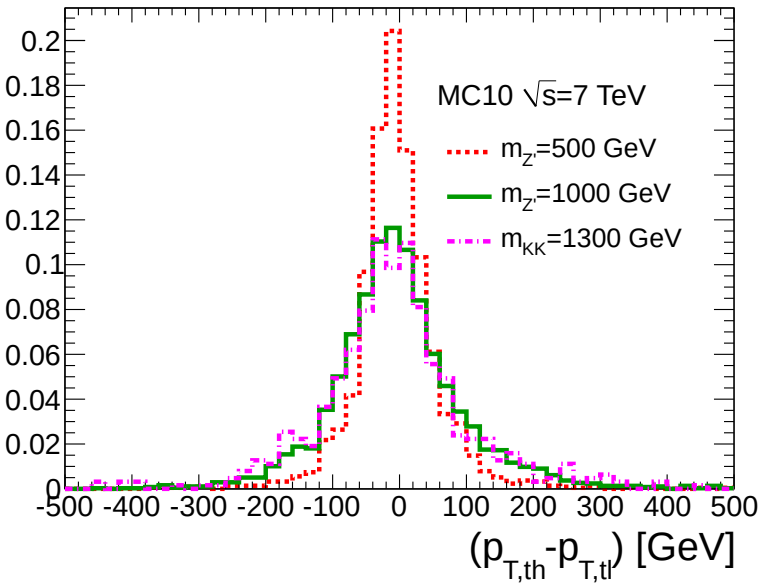


Figure 6.17: Difference of transverse momentum between the hadronic and leptonic top calculated using matched events for three different resonances masses: $m_{Z'} = 500$, 1000 and $m_{KK} = 1300$ GeV.

6.14c, respectively. A better correlation between the reconstructed $t\bar{t}$ mass and the true mass for the SM $t\bar{t}$ sample is observed. However, for higher masses a large left tail is observed in Figure 6.12f indicating an underestimation of $m_{t\bar{t}}$. The reconstruction efficiency achieved using the χ^2 algorithm is the highest one between the three reconstruction algorithms described so far, as shown in Figure 6.15.

Summarizing, the χ^2 algorithm reconstruction efficiency is higher than the one achieved by the four hardest jets and dRmin algorithm. However, its resolution is the worst between the three algorithms. To try to understand this behavior two different studies were done. First, the $m_{t\bar{t}}$ distribution was plotted for matched events and for non matched events in Figure 6.18. dRmin is able to find a good approximation of the mass, even for non matched events. χ^2 seems to be more sensitive to the ISR/FSR and pile-up effects. To verify this statement, the pile-up dependence of the reconstruction efficiency was studied. Figures 6.19 and 6.20 show that the reconstruction efficiency for the four hardest jets and dRmin algorithms remains stable as a function of the average number pp collisions per bunch crossing, $< \mu >$, and the number of reconstructed primary vertices. On the other hand, the reconstruction efficiency for the χ^2 decrease when the number of pile-up events increases.

6.9.5 Scaled χ^2 algorithm

The scaled χ^2 algorithm is a variant of the χ^2 algorithm. The χ^2 algorithm tries to identify all the reconstructed objects in the final state of the $t\bar{t}$ pair. It uses reconstructed jets calibrated at the EM+JES scale. The EM+JES calibration as discussed in Section 4.7.4.1 corrects the energy and momentum of the calorimeter jets, using the kinematics of the corresponding Monte Carlo particle jet as a reference. The objective of the scaled χ^2 algorithm is to calibrate jets to their partonic scale instead of the particle jet scale, in order to improve the resolution of the $t\bar{t}$ invariant mass. For events without a jet with $m^{\text{jet}} > 60$ GeV, the light jets selected by the χ^2 method as coming from the W boson decay are scaled to the W boson mass, M_W, which is

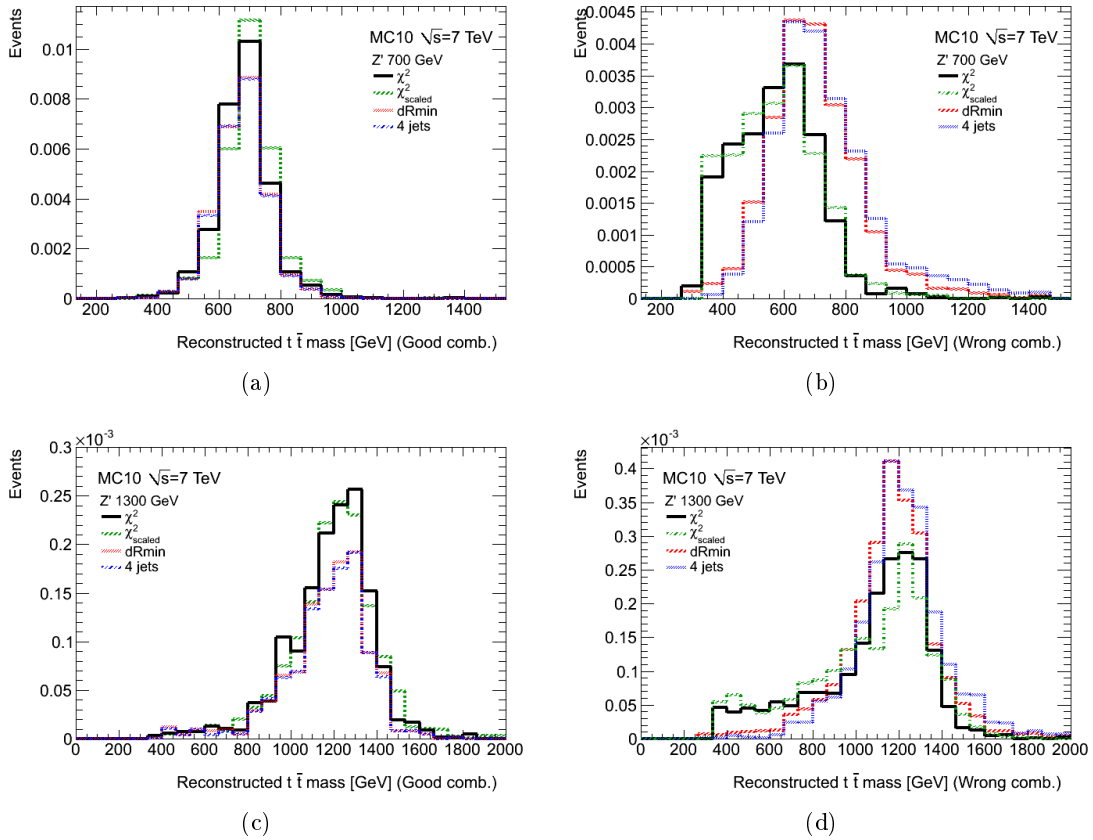


Figure 6.18: $m_{t\bar{t}}$ distribution for matched events (left) and for non matched events (right) for Z' with $m_{Z'} = 700$ GeV (up) and $m_{Z'} = 1300$ GeV (bottom).

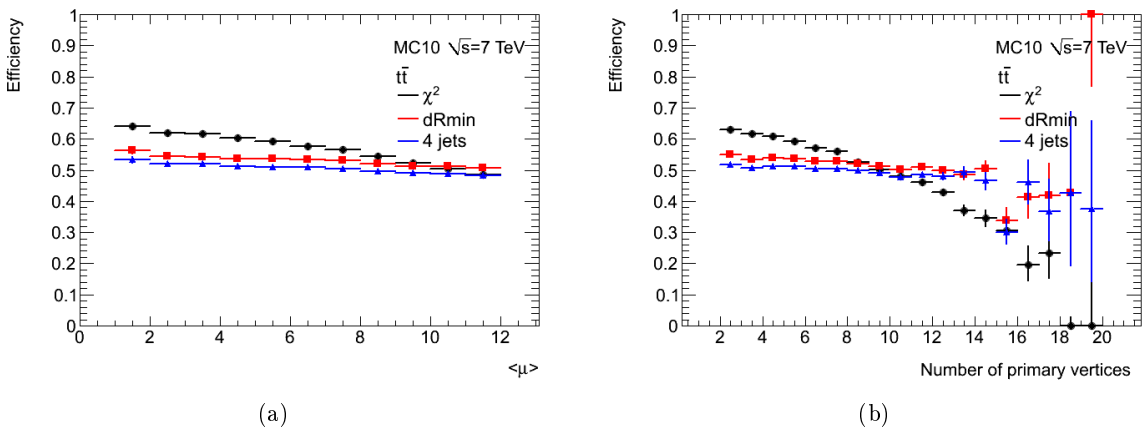


Figure 6.19: Reconstruction efficiency using the four hardest jets, dRmin, χ^2 (same for χ^2 scaled) methods as a function of (a) the average number pp collisions per bunch crossing, $< \mu >$ (left) and (b) the number of reconstructed primary vertices (right) for the SM $t\bar{t}$ sample.

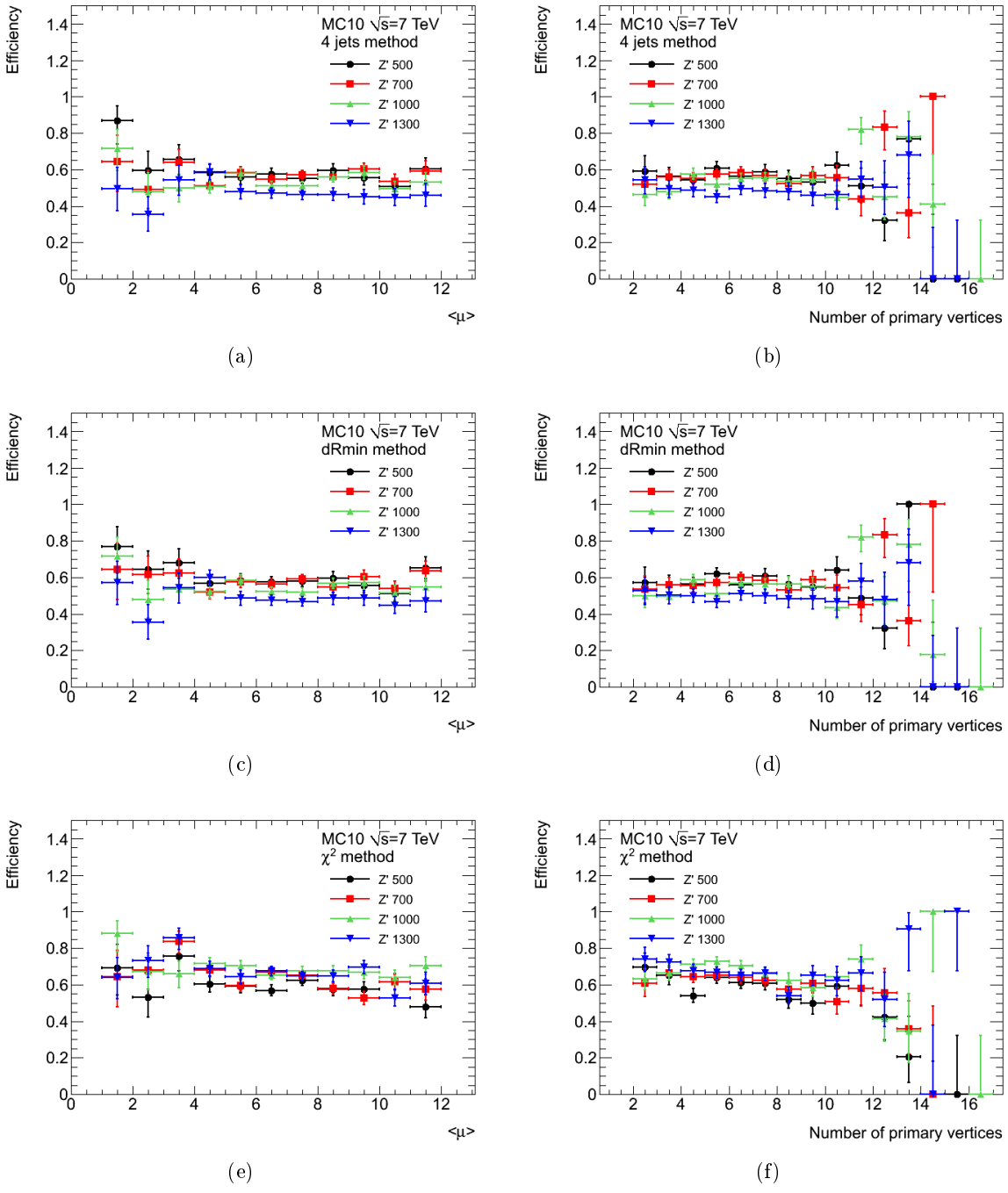


Figure 6.20: Reconstruction efficiency using the (a,b) four hardest jets, (c,d) dRmin and (e,f) χ^2 method as a function of the average number pp collisions per bunch crossing, $< \mu >$ (left) and the number of reconstructed primary vertices (right) for different Z' samples.

known with a high precision (of the order of 30 MeV [13]). The jets associated to the $b-$jets produced in the top-quark decay are scaled to the top-quark mass, M_{top}. For events with at least one jet with $m^{\text{jet}} > 60$ GeV where two of the reconstructed are assumed to be merged, the scaling to the W boson mass is not done. But the two jets used to reconstruct the hadronic top quark are rescaled to the top-quark mass. Three different scalings are applied:

6.9.5.1 Hadronic W mass scaling

For events without a jet with $m^{\text{jet}} > 60$ GeV, the two jets associated to the hadronic W boson decay are rescaled to the PDG value of the W boson mass, $M_W = 80.4$ GeV. The scaling factor is given by:

$$\alpha = \frac{M_W}{m_W^{nocal}}, \tag{6.8}$$

where m_W^{nocal} corresponds to the W boson invariant mass calculated using the two chosen jets. No scaling is applied in events with a jet with $m^{\text{jet}} > 60$ GeV, since the χ^2 method assume that the merging can not only occurs between both quarks from W boson decay, but also between one quark from W boson decay and the $b-$quark from the top-quark decay.

6.9.5.2 Top-quark mass scaling

For events without a jet with $m^{\text{jet}} > 60$ GeV, the chosen jet associated to the $b-$quark from the top quark decay (either leptonic or hadronic) is rescaled to the PDG value of the top-quark mass, $M_{top} = 172.5$ GeV. The scaling factor corresponds to the positive solution of the quadratic equation $a\beta^2 + b\beta + c$ where:

$$a = m_b^2 \tag{6.9}$$
$$b = m_{top}^2 - m_W^2 - m_b^2 \tag{6.10}$$
$$c = -M_{top}^2 + m_W^2, \tag{6.11}$$

where m_b is the invariant mass of the jet associated to the $b-$quark and m_{top} is the top-quark invariant mass calculated using the three jets chosen by the χ^2 algorithm. In the case of the hadronic top quark, m_W is used after the W mass scaling. For events with a jet with $m^{\text{jet}} > 60$ GeV, the jet associated to the $b-$quark in the leptonic side is scaled by β. In the case of the hadronic top-quark decay the situation is different since the two jets chosen by the χ^2 method to built the hadronic top quark will be directly rescaled to the PDG value of the top-quark mass using a scaling factor given by:

$$\gamma = \frac{M_{top}}{m_{top}^{nocal}}, \tag{6.12}$$

where m_{top}^{nocal} corresponds to the top-quark invariant mass calculated using the two chosen jets. Figure 6.21 and 6.22 show the reconstructed $t\bar{t}$ mass resolution and the reconstructed $t\bar{t}$ mass spectra before and after the different rescalings for the SM $t\bar{t}$ samples and three different Z' resonances masses, using only matchable events. For matchable events and SM $t\bar{t}$ and low resonance mass samples the reconstructed $t\bar{t}$ mass resolution is improved after applying the χ^2 scaled method. The method is built to calibrate the jets produced in the top-quark decay, but when the jets selected by the χ^2 method do not come from the top-quark decay the scaling introduces left tails in the reconstructed $t\bar{t}$ mass. This is the case of high resonance masses as is shown in Figure 6.22. The effect is more pronounced when we look at the reconstructed mass and the corresponding mass resolution for all events in Figures 6.12e and 6.12h. The reconstructed

$t\bar{t}$ mass is shown as a function of true mass for SM $t\bar{t}$ production and Z' with $m_{Z'} = 2000$ GeV in Figures 6.13d and 6.14d, respectively. All events were considered.

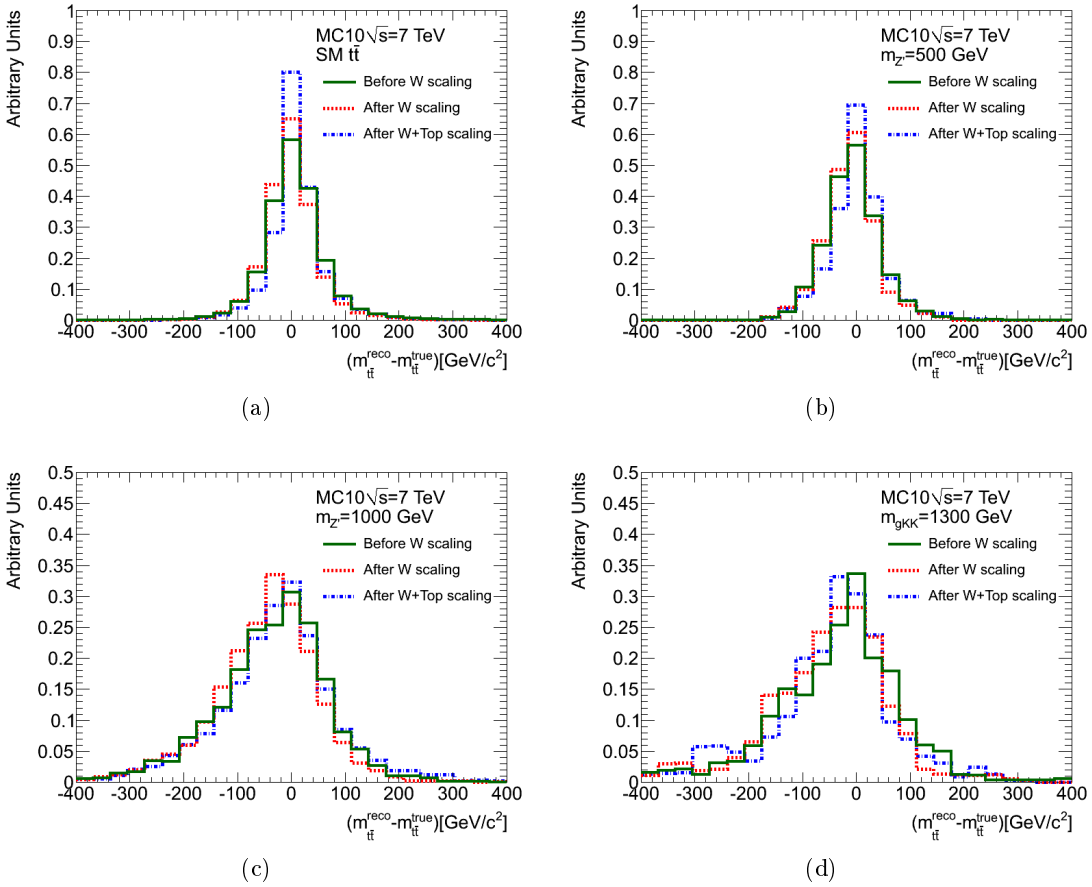


Figure 6.21: Reconstructed $m_{t\bar{t}}$ resolution for matchable events before and after χ^2 scalings for (a) SM $t\bar{t}$, (b) Z' $m_{Z'} = 500$ GeV, (c) Z' $m_{Z'} = 1000$ GeV and (d) g_{KK} $m_{g_{KK}} = 1300$ GeV.

Figure 6.23 shows the agreement between data and expectation from the sum of all backgrounds in the $t\bar{t}$ mass spectra. Both channels have been combined. Two signal points are also shown, a Z' with $m_{Z'} = 800$ GeV and a KK gluon with $m_{g_{KK}} = 1300$ GeV. Figure 6.24 shows the relative difference in reconstructed $t\bar{t}$ mass between data and expectation for the different methods. A data deficit is observed when using the χ^2 method around 1 TeV.

The χ^2 method is sensitive to the pile-up and the efficiency of the four hardest jets method is smaller than the one obtained when using the dRmin method. Therefore, the dRmin method has been chosen as the reconstruction method to be used in the analysis. The impact of pile-up on the reconstructed $t\bar{t}$ mass for different samples is shown Figure 6.26. Small differences are seen between high and low pile-up events.

6.10 Systematic uncertainties

This analysis is affected by several sources of systematic uncertainties. They can affect only the normalization of the signal and background events or affect the normalization and the shape of the reconstructed $t\bar{t}$ mass distribution. The impact of the latter depends on the mass reconstruction method used. As mentioned in the previous section, the dRmin method has been chosen as the default $t\bar{t}$ reconstruction method.

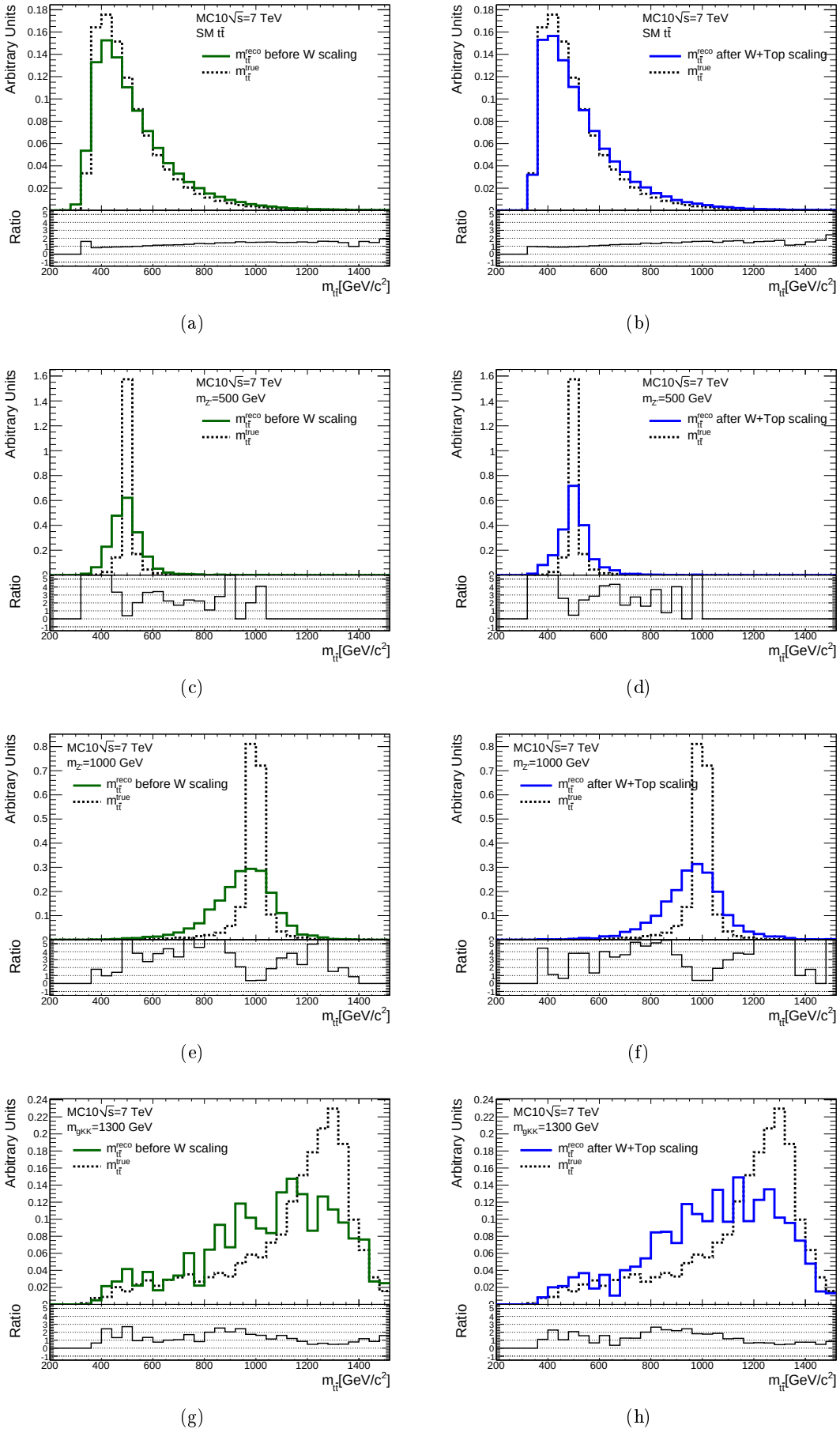


Figure 6.22: Reconstructed $m_{t\bar{t}}$ for matchable events before and after χ^2 scalings for (a,b) SM $t\bar{t}$, (c,d) Z' $m_{Z'} = 500$ GeV, (e,f) Z' $m_{Z'} = 1000$ GeV and (g,h) g_{KK} $m_{g_{KK}} = 1300$ GeV. The generated $t\bar{t}$ mass, $m_{t\bar{t}}^{true}$ is shown for reference.

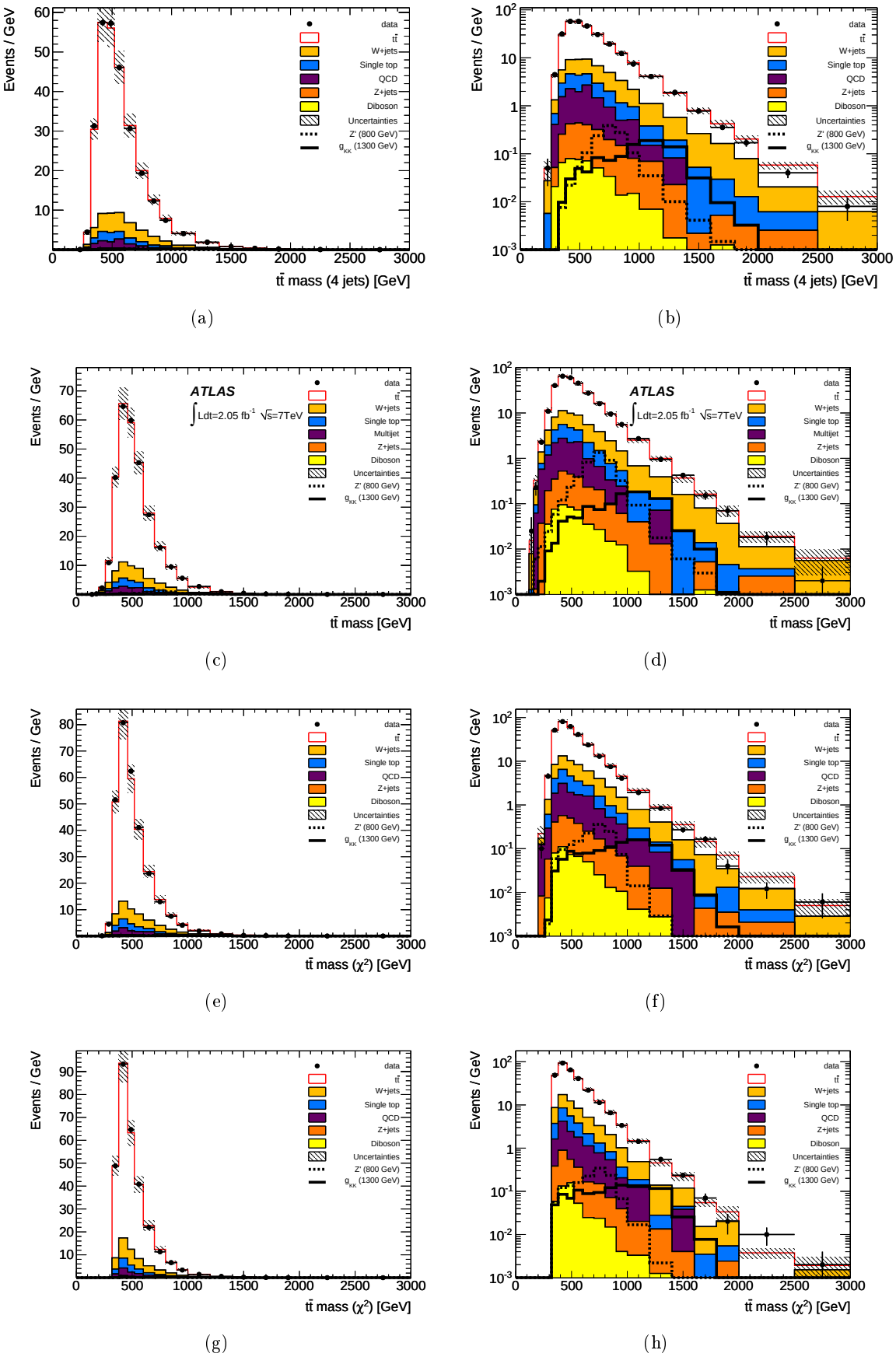


Figure 6.23: Reconstructed $t\bar{t}$ mass using the (a,b) four hardest jets, (c,d) dRmin, (e,f) χ^2 and (g,h) χ^2 scaled methods in linear (left) and logarithmic scales (right). The grey hashed area shows the total background normalization uncertainty. Both channels have been combined.

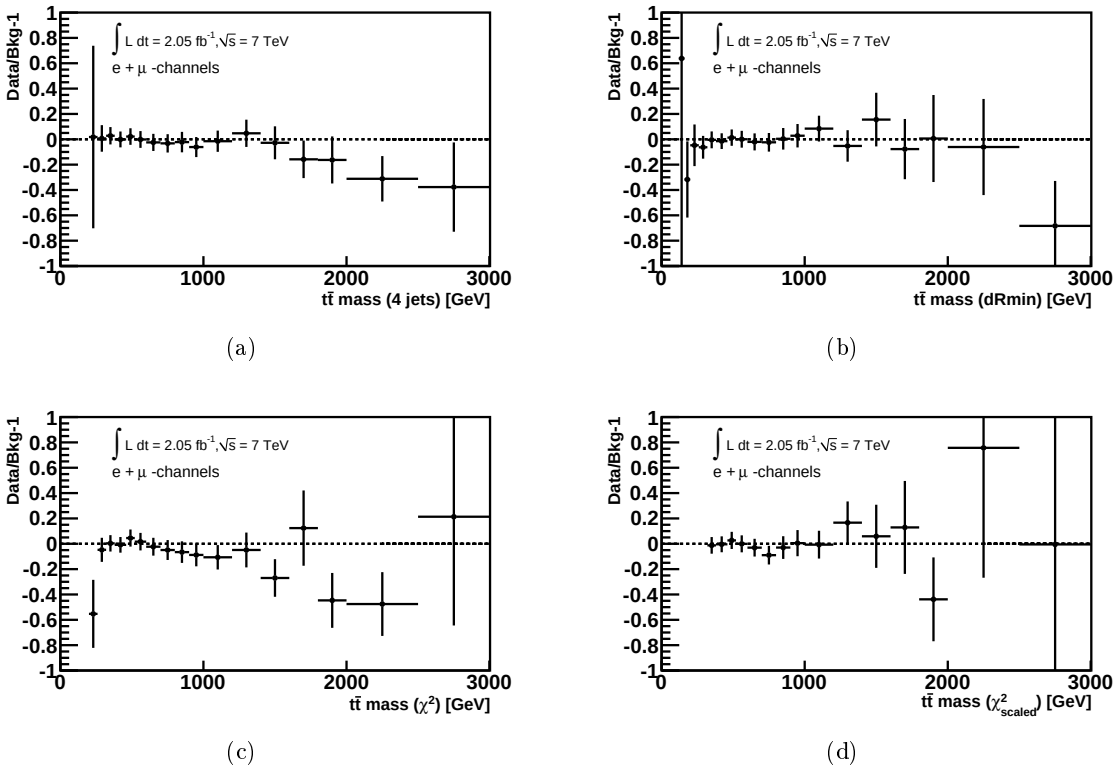


Figure 6.24: Relative difference in reconstructed $t\bar{t}$ mass between data and expectation for the (a) four hardest jets, (b) dRmin, (c) χ^2 and (d) χ^2 scaled methods.

6.10.1 Systematic uncertainties affecting the normalization only

Luminosity The estimated integrated luminosity is determined using ATLAS subdetectors, as described in Section 2.2.5. Its associated uncertainty was calculated to be 3.7% [56]. This uncertainty is applied to all Monte Carlo samples, except the W+jets samples since their normalization is estimated using data (see Section 6.7.2).

Theoretical cross section Normalizations of the $t\bar{t}$ ($^{+7.0}_{-9.6}$% [159]), single top (10%), Z+jets (48%) and diboson (5%) backgrounds [160] are varied within their uncertainties. The normalization uncertainty due to the PDF choice is included in the overall $t\bar{t}$ normalization uncertainty quoted above. The PDF uncertainty for signal is illustrated as a band around the predicted cross-section.

Multijet normalization An uncertainty of 50% is applied to the multijet background rate in both channels (see Section 6.7.3)

W+jets normalization The uncertainty associated to the W+jets background rate corresponds to 48% as described in Section 6.7.2.

Electron trigger and reconstruction uncertainty The uncertainties on the electron trigger and reconstruction scale factors correspond to a flat 1% and 1.5% uncertainty, respectively (see Section 4.5.3).

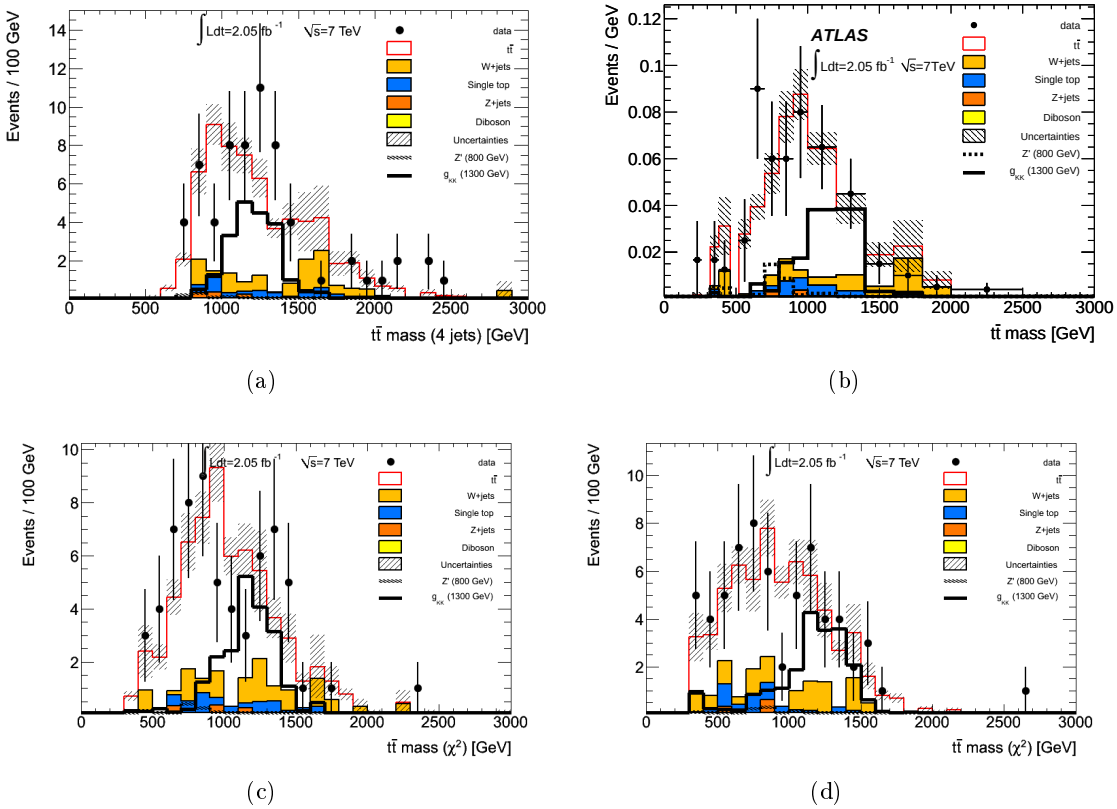


Figure 6.25: Reconstructed $t\bar{t}$ mass for events with at least one jet with $m^{\text{jet}} > 60$ GeV using the (a,b) four hardest jets, (c,d) dRmin, (e,f) χ^2 and (g,h) χ^2 scaled methods in linear scale. The grey hashed area shows the total background normalization uncertainty. Both channels have been combined.

6.10.2 Systematic uncertainties affecting the normalization and $m_{t\bar{t}}$ shape

Electrons related uncertainties There are three uncertainties associated to the electrons that affect the $m_{t\bar{t}}$ shape: the electron identification, the electron energy scale and the electron energy resolution uncertainties. The electron identification uncertainty corresponds to the uncertainty on the electron identification scale factors (see Section 4.5.3). It varies with η_{cluster} and E_{T} between 3% and 5% and is applied in an event by event basis. In order to take into account discrepancies between the electron energy resolution on simulation and data, the electron energy for simulation events is smeared. In addition, the electron energy in data was corrected to match the Z boson peak mass. Electron energy scale and resolution uncertainties of the order of 1% to 2% are applied to the simulated samples in an object by object basis.

Muons related uncertainties The uncertainties on the corresponding muon trigger, identification and reconstruction scale factors are applied (see Section 4.6.2). These uncertainties are of the order of 1%-1.5%. In release 16 there was an issue with the trigger information in Monte Carlo and the trigger matching requirement has not been applied. An additional 1.5% uncertainty is applied to account for this. Muon momentum resolution and energy scales are varied up and down within uncertainties as prescribed by the ATLAS Muon Combined Performance Group, taking into account properly the differences between the corrections applied to ID and MS separately.

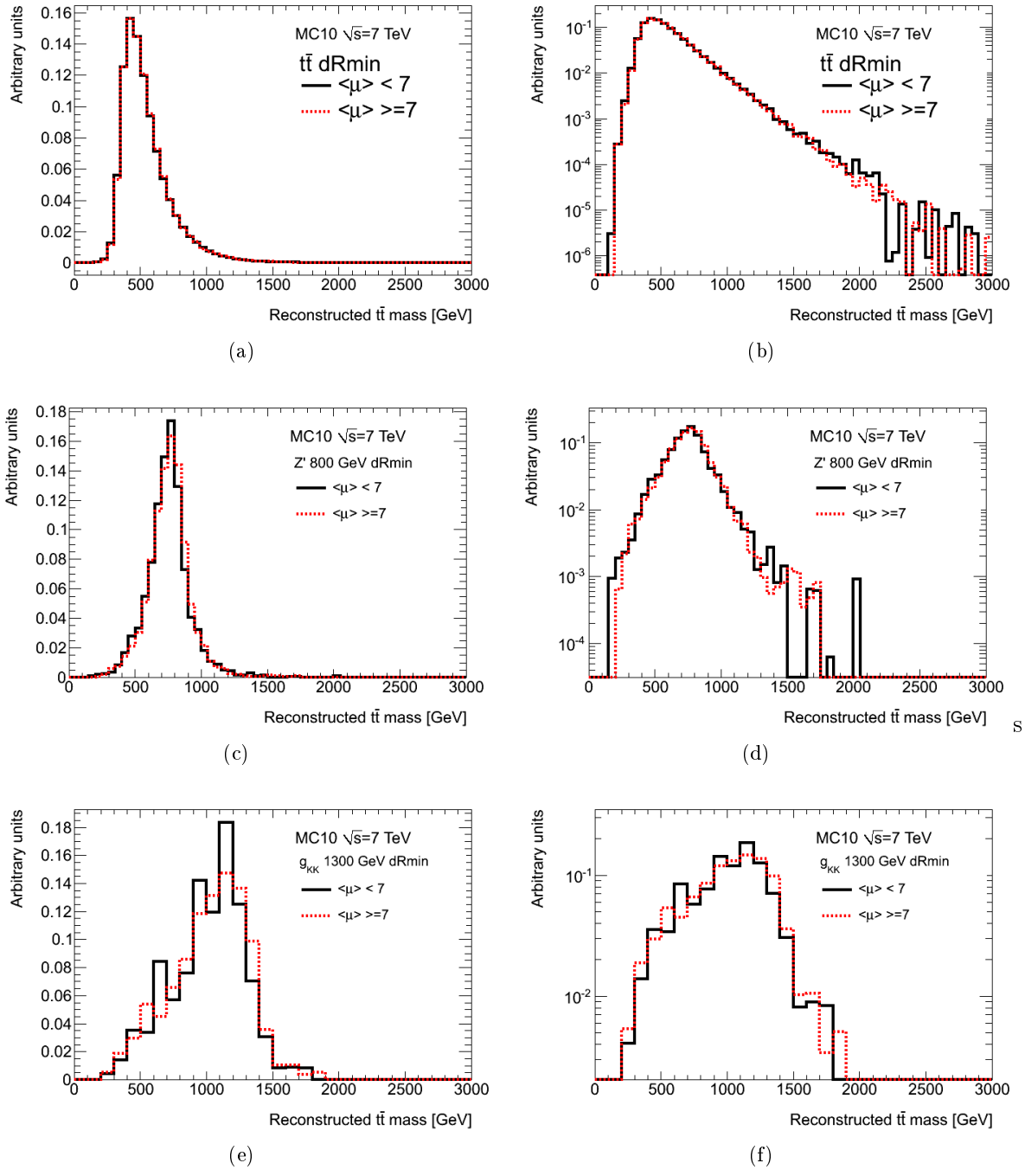


Figure 6.26: Impact of pile-up on the reconstructed $t\bar{t}$ mass for the (a,b) SM $t\bar{t}$ background, the (c,d) Z' $m_{Z'} = 800$ GeV and the (e,f) g_{KK} $m_{g_{KK}} = 1300$ GeV sample in linear (left) and logarithmic (right) scale.

Jets-related uncertainties There are five uncertainties associated with jets: the jet energy scale (JES), the pile up, the $b-$jet energy scale, the jet energy resolution (JER) and the jet reconstruction efficiency uncertainties. The JES uncertainty is the systematic uncertainty on the EM+JES calibration resulting from uncertainties on calorimeter response, dead material description, calorimeter cell noise thresholds description, fragmentation and underlying event modeling, the topology and flavour composition [3, 116]. Without taking into account the flavor composition contribution and for $|\eta| < 2.5$, the JES uncertainty varies from about 4% for high p_{T} jets to about 9% for low p_{T} jets. The quark/gluon flavor composition contribution varies with the samples used. A pile-up uncertainty on the JES is also applied to account for differences in the jet energy scale in pile-up conditions different from that simulated in the sample used to derive the EM+JES calibration. It varies from about 2% for high p_{T} jets to about 5% for low p_{T} jets. The $b-$jet energy scale uncertainty accounts for the difference of the energy scale for $b-$jets with respect to light quark jets. $b-$jets are scaled by a factor that varies from about 0.8% in the high p_{T} region to about 2.5% in the low p_{T} region. The pile-up and $b-$jet uncertainties are added in quadrature to the JES uncertainty. The JER uncertainty accounts for differences in the resolution on the jet energy in data and in Monte Carlo [109]. It is y and p_{T}-dependent. It is evaluated by oversmearing the jet energy up to the JER uncertainty and then it is symmetrized. Finally the jet reconstruction efficiency uncertainty is evaluated by randomly dropping jets from events with a p_{T} and $|\eta|$-dependent probability of about 2% [109].

$b-$tagging uncertainty The uncertainties on the $b-$tagging, $c-$tagging, $\tau-$tagging and mistagging scale factors determined by the Flavor Tagging Performance Group are applied. In release 16 the $c-$tagging efficiency is considered to be fully correlated with the b-tagging efficiency.

$E_{\mathrm{T}}^{\mathrm{miss}}$ Any variation in an object's p_x or p_y is propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement. Two $E_{\mathrm{T}}^{\mathrm{miss}}$ related systematic uncertainties are applied [122]. The first accounts for the energy scale and energy resolution effects propagated to the MET. In this case an uncertainty is applied on the energy scale of energy deposits not associated with any objects ("cellout") and on the soft jets. The second one accounts for the pile-up effects. The uncertainty on $E_{\mathrm{T}}^{\mathrm{miss}}$ due to pile-up effects is a flat 10% uncertainty on the contributions not associated with any objects.

LAr hole In data recorded during the period during which some of the front-end boards in LAr calorimeter could not be read out, events affected by the LAr hole are vetoed. Having occurred after the production of the used Monte Carlo samples, it was necessary to apply the same procedure to simulated events to correct the mismatch between data and simulated events. A systematic uncertainty corresponding to varying the jet energy threshold for jets considered as potentially being in the LAr hole by 4 GeV is applied.

$t\bar{t}$ background related uncertainties There are four sources of systematic uncertainties affecting the SM $t\bar{t}$ background that have been studied: the modeling of the hard process, the simulation of the ISR and FSR, the modeling of the parton shower and fragmentation and the PDF choice.

The systematic uncertainty associated to the modeling of the hard process is addressed by comparing the $t\bar{t}$ mass spectra using two different Monte Carlo generators: MC@NLO and POWHEG.

The impact of the ISR and FSR modelization is estimated by comparing the $t\bar{t}$ mass spectra between the nominal MC@NLO sample and samples generated with ACERMC

and showered with PYTHIA where the ISR and FSR parameters were varied in order to produce less or more gluon radiation. Six samples were generated: more ISR, less ISR, more FSR, less ISR, more ISR and FSR and less ISR and FSR. The sample with the largest deviation from the nominal one is used, which in this case corresponds to the varied ISR only.

The impact of the parton shower and fragmentation model is estimated by a comparison of POWHEG+HERWIG with POWHEG+PYTHIA samples. These three uncertainties are determined as a function of generated $t\bar{t}$ mass and applied to the baseline $t\bar{t}$ and signal samples.

Finally, the difference in shape between the predictions obtained using `CTEQ66` and `MSTW2008nlo` is used to estimate the PDF shape uncertainty. The dependence of the reconstructed $m_{t\bar{t}}$ spectra for different Monte Carlo samples generated with different values of the top-quark mass was also studied. The results shown a small change in yield ($+1.1\%, -2.0\%$ for a 1 GeV variation in m_t) and no significant change in the spectrum within available statistics for $m_{t\bar{t}} > 400$ GeV. Therefore, this uncertainty is not considered in the limit setting.

W+jets background related uncertainties The W+jets shape uncertainty is determined by modifying some generator parameters such as the function which gives the factorization scale for the PDF (iqopt) and the minimum p_{T} to consider a parton as a hard parton (ptjmin) with respect to the nominal sample (see Section 3.2.2.3). The flavor composition of the W+jets background is varied as described in Section 6.7.2.

Multijet background related uncertainties The shape uncertainty on the multijet background is described in Section 6.7.3.

The impact of each systematic uncertainties affecting the normalization and $m_{t\bar{t}}$ shape on the total expected background yield and for the Z' samples with $m_{Z'} = 1$ TeV is presented in Table 6.9. The dominant sources of shape systematic uncertainties are the $b-$tagging efficiency, the JES, the modeling of ISR and FSR and the $Wc/Wbb/Wcc$ normalization. The impact of some of the shape systematic uncertainties in the reconstructed $m_{t\bar{t}}$ mass spectra is shown in Figures D.6 and D.7 and Figures E.6 and E.7 for electron and muon channel, respectively.

6.11 Results

The presence of a new heavy resonance decaying into a $t\bar{t}$ pair can be observed as a bump or data excess in the reconstructed $t\bar{t}$ mass spectra. The reconstructed $t\bar{t}$ mass spectra obtained in Section 6.9 can be compared with the background-only and signal-plus-background hypothesis. If it is compatible with the signal-plus-background hypothesis a possible discovery could be claimed. Otherwise, if it is compatible with the background-only hypothesis, an upper limit on the signal cross section time branching ratio, $\sigma_{Z'/g_{KK}} \times BR(Z'/g_{KK} \to t\bar{t})$, can be set.

To compare the data reconstructed $t\bar{t}$ mass spectra with the background-only hypothesis the BumpHunter [161] algorithm is used. The BumpHunter algorithm performs a statistical test to evaluate if the data is consistent with a given hypothesis (the background-only hypothesis in this case). The statistical test returns a quantity known as the $p-$value, which corresponds to the probability of observing in pseudo-experiments a deviation at least as signal-like as the one observed in data. The $p-$value can be transformed into a standard deviation significance, σ, with negative values indicating the absence of any signal. After running the BumpHunter no significant deviations (data excesses or deficits) from the SM expectations are observed. Table 6.10

Source	$t\bar{t}$	W+jets	All MC	Z', $m_{Z'} = 1000$ GeV
Jet energy scale	+7.7%	+58.6%	+14.5%	+3.4%
	-8.2%	-32.2%	-11.6%	-4.2%
Jet energy resolution	±1.9%	±14.5%	±3.9%	±0.2%
Jet reconstruction efficiency	< 1%	< 1%	< 1%	< 1%
b-tagging efficiency	+15.1%	+5.4%	+12.7%	+18.9%
(incl. mistag rate)	-13.4%	-4.8%	-11.2%	-16.5%
Muon p_{T} resolution	< 1%	< 1%	< 1%	< 1%
Muon efficiency	±1.1%	±1%	±1.1%	±1%
Electron energy scale	< 1%	< 1%	< 1%	< 1%
Electron energy resolution	< 1%	< 1%	< 1%	< 1%
Electron ID efficiency	±1.1%	±1.1%	±1.1%	±1.2%
$E_{\mathrm{T}}^{\mathrm{miss}}$	< 1%	< 1%	< 1%	< 1%
$E_{\mathrm{T}}^{\mathrm{miss}}$ Pile-Up	< 1%	< 1%	< 1%	< 1%
$m_{t\bar{t}}$ shape (MC generator)	<1%	–	–	–
Parton shower & fragmentation	±4.1%	–	–	–
ISR	+4.3%	–	–	+0.9%
	-10.4%	–	–	-8.5%
$t\bar{t}$ PDF	±2.2%	–	–	–
Wc normalization	–	± 3.6%	–	–
Wbb/Wcc normalization	–	± 20.1%	–	–
W modeling: iqopt3	–	± 3.8%	–	–
W modeling: ptjmin	–	< 1%	–	–
LAr hole	< 1%	< 1%	< 1%	< 1%

Table 6.9: Impact of the shape systematic uncertainties on the background event yields and for the Z' samples with $m_{Z'} = 1$ TeV after all selection cuts.

channel	p-value	σ	Mass range (GeV)
Excess Search			
Electron	0.715 ± 0.005	-0.57	1280 – 1800
Muon	0.185 ± 0.004	0.90	400 – 640
Added	0.319 ± 0.005	0.47	920 – 1600
Deficit Search			
Electron	0.119 ± 0.003	1.18	240 – 640
Muon	0.691 ± 0.005	-0.50	720 – 920
Added	0.551 ± 0.005	-0.13	160 – 320

Table 6.10: BumpHunter results for the search of deviations in the reconstructed $t\bar{t}$ mass spectra. The "Added" rows correspond to adding the electron and muon histograms bin-by-bin before running the BumpHunter[8].

shows the results from the BumpHunter algorithm. Systematic uncertainties were taken into account.

Given the absence of a signal, upper limits were set on $\sigma_{Z'/g_{KK}} \times BR(Z'/g_{KK} \to t\bar{t})$ as

a function of the signal mass. A bayesian approach was used [162, 163]. The probability of observing a count D (observed number of data events), if the mean count was d (signal plus background expectation) in a given bin of the reconstruction $t\bar{t}$ mass distribution is given by a Poisson distribution:

$$p(D|d) = \frac{e^{-d}d^{D}}{D!}. \tag{6.13}$$

The signal expectation is given by the product of the signal acceptance, α, the signal cross section, σ, and the integrated luminosity, $\mathtt{L}$. The product of α and $\mathtt{L}$ is known as the effective luminosity, a. The background expectation, b, is given by the sum of the yield of all the background sources in the analysis. As the probability of observing a count in a given bin is independent of the counts in the others bin, the likelihood of the distribution of counts is given by the product of the Poisson distributions over all bins of the reconstructed $t\bar{t}$ invariant mass.

$$L(\mathbf{D}|\mathbf{d}) = L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b}) = \prod_{i}^{n} p(D_i|d_i), \tag{6.14}$$

where $\mathbf{D}$, $\mathbf{d}$, $\mathbf{a}$ and $\mathbf{b}$ correspond to the vectors of D, d, a and b in n bins. To calculate a likelihood for combined channels, the likehoods of the individual channels are multiplied. A variable-size binning is used in the reconstructed $t\bar{t}$ mass spectra, with bins ranging in size from 40 to 500 GeV bins for narrow resonances, and 80 to 500 GeV for KK gluons. These values are close to the mass resolution while limiting bin-by-bin statistical fluctuations. A single bin contains all events with $m_{t\bar{t}} > 2.5$ TeV. The region $m_{t\bar{t}} < 500$ GeV (SM $t\bar{t}$ threshold region) is not considered. The posterior probability density function for the signal cross section, $p(\sigma|\mathbf{D})$, is calculated using the Bayes' theorem. $p(\sigma|\mathbf{D})$ represents the probability that σ have a certain value given the observed data yields, $\mathbf{D}$:

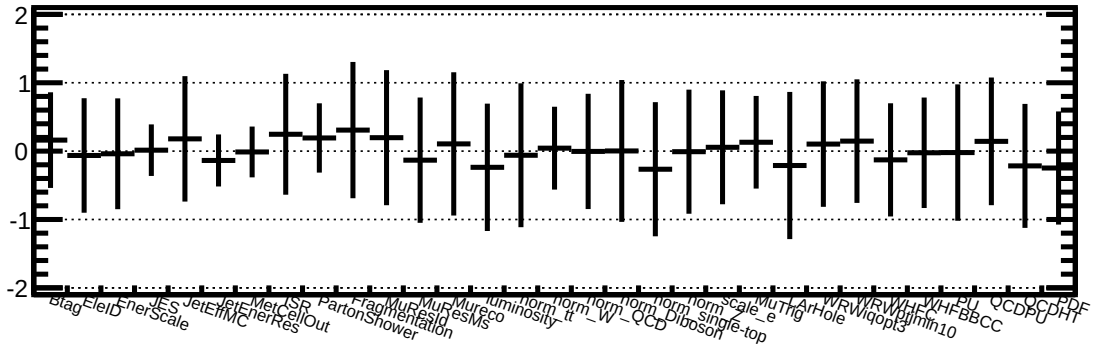
$$p(\sigma|\mathbf{D}) = \frac{1}{N}\int\int L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b})\pi(\mathbf{a}, \mathbf{b})\pi(\sigma)d\mathbf{a}\ d\mathbf{b}, \tag{6.15}$$

where N is the overall normalization, $\pi(\sigma)$ is the prior probability density in the signal cross section, which quantifies our knowledge for σ and $\pi(\mathbf{a}, \mathbf{b})$ encodes the knowledge of $\mathbf{a}$ and $\mathbf{b}$. The $\pi(\sigma)$ used was flat and positive.

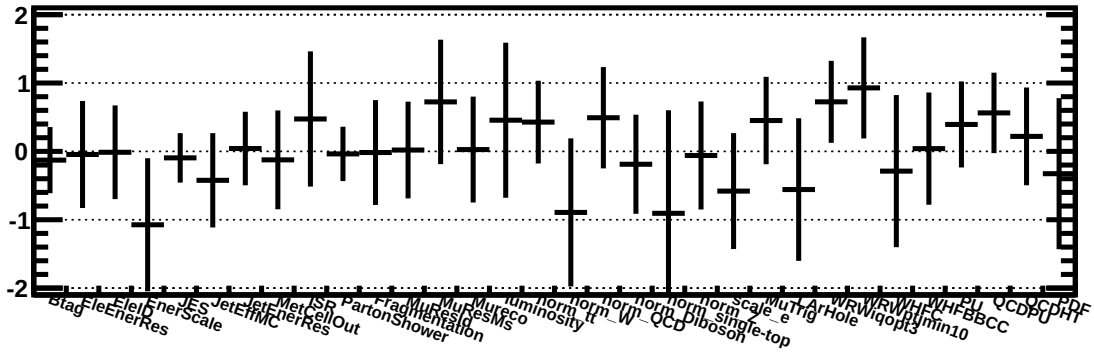
Systematic uncertainties are incorporated by smearing the parameters of the poisson distribution in each bin. For each single bin, the bin content is Poisson fluctuated 1000000 times. This procedure is known as a pseudo-experiment. For each systematic uncertainty a gaussian prior (which can be directly associated with the prior $\pi(\mathbf{a}, \mathbf{b})$) controls the probability for a given deviation of the parameter from the nominal value. The gaussian prior has mean 0 and width of 1. A likelihood is built for each pseudo-experiment, and the posterior density distribution is obtained by summing up all the individual likelihoods. Because the likelihood curves are added, combinations of systematic uncertainties shifts that lead to good agreement get a larger weight than combinations that lead to disagreements. The upper observed limit, σ_{upper}, on $\sigma_{Z'/g_{KK}} \times BR(Z'/g_{KK} \to t\bar{t})$ is identified with the 95% point of the posterior density function:

$$\int_{0}^{\sigma_{\text{upper}}} p(\sigma|\mathbf{D}) = 95\%. \tag{6.16}$$

The expected limits are obtained using the same procedure, but using setting the count D in every bin to the expected background yield in that bin. Around 5000 pseudo-experiments are done to get a distribution of limits. A large number of samples of systematic shifts are created (about 5000000 for the observed limits and 20000 for the expected limits) for each pseudo-experiment. The one and two sigma bands around the expected limits correspond to the one



(a)



(b)

Figure 6.27: Mean and width (error bars) of the posterior distribution for each systematic uncertainty for (a) expected and (b) observed limits. The test signal used for the plot is a Z' boson with $m_{Z'} = 1$ TeV. The bins labeled "Fragmentation" and "Parton Shower" correspond to the $m_{t\bar{t}}$ shape and fragmentation and parton shower systematic uncertainties, respectively [8].

and two sigma uncertainty on the expected limit and are determined from the distribution of upper limits in pseudo-experiments. Figure 6.27 shows the mean and width of the posterior distribution for each systematic uncertainty for expected and observed limits for a Z' boson with $m_{Z'} = 1$ TeV. As the prior used was a gaussian with mean 0 and width 1, then the posterior distribution should also have a mean 0 and width 1 for perfectly estimated central values and systematic uncertainties. Figure 6.27a shows that the reconstructed $t\bar{t}$ invariant mass distribution has the ability to constrain the impact of some of the systematic uncertainties on this distribution (as is the case for the JES and JER systematic uncertainties). Since the data is not in perfect agreement with the expectation, the posteriors for the observed limits show means which are sometimes slightly different from zero, as seen in Figure 6.27b.

The resulting limits can be seen in Figure 6.28 and Table 6.11. When including the systematic uncertainties the observed (expected) limits on $\sigma\times$ BR($Z' \to t\bar{t}$) ranges from 9.3 (8.5) pb at $m_{Z'} = 500$ GeV to 0.95 (0.62) pb at $m_{Z'} = 1300$ GeV, excluding 500 GeV $< m_{Z'} <$ 880 GeV at 95% C.L. The expected mass exclusion is 500 GeV $< m_{Z'} <$ 1010 GeV at 95% C.L. When using the old LO cross section for the Z' samples (see Section 6.6) the observed (expected) exclusion limit is 500 GeV $< m_{Z'} <$ 860 (930) GeV (for comparison with Tevatron results). The observed (expected) limits on $\sigma\times$ BR($g_{KK} \to t\bar{t}$) ranges from 10.1 (10.3) pb at $m_{g_{KK}} = 500$ GeV to 1.6 (0.9) pb at $m_{g_{KK}} = 1300$ GeV, excluding resonances with mass between 500 and 1130 GeV at 95% C.L. The expected mass exclusion is 500 GeV $< m_{g_{KK}} <$ 1360 GeV at 95% C.L.

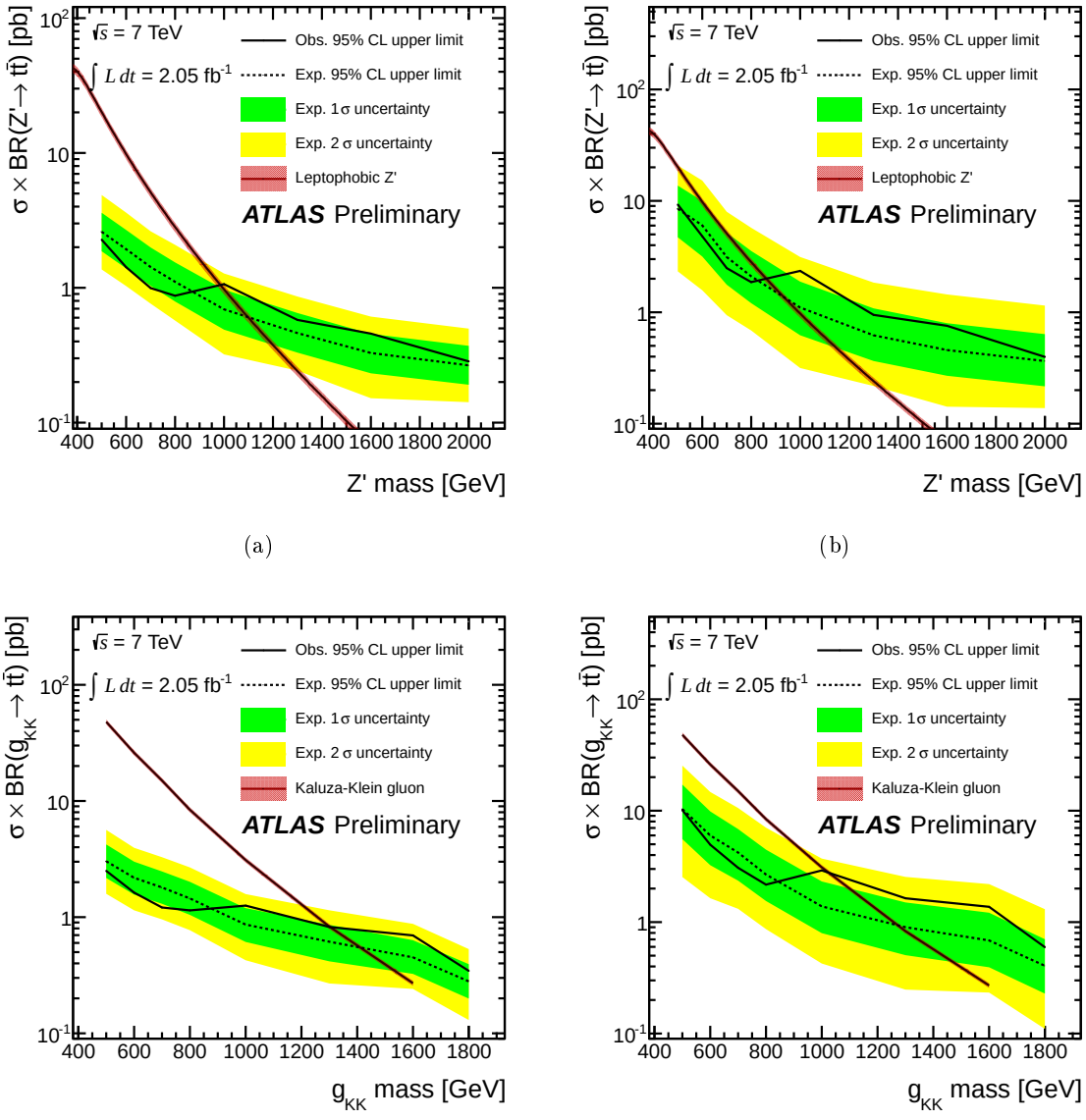


Figure 6.28: Expected (dashed line) and observed (line) upper limits on $\sigma\times$ BR($Z' \to t\bar{t}$) (top) and $\sigma\times$ BR($g_{KK} \to t\bar{t}$) (bottom) using the dRmin mass reconstruction method. Limits with (right) and without (left) consideration of systematic uncertainties are shown. The green and yellow bands show the range in which the limit is expected to be in 68% and 95% of experiments, respectively, and the red lines correspond to the predicted signal cross-sections. The red band around the topcolor cross-section represent the PDF uncertainty [8].

6.12 Summary

In this chapter the search for new resonances decaying to top quark pairs in the lepton plus jets channel has been presented. The search has been done using the $t\bar{t}$ invariant mass spectra as the discriminant variable and using 2.05 fb^{-1} of data collected with the ATLAS detector during 2011 in pp collisions at $\sqrt{s} = 7$ TeV. The lepton plus jets channel is characterized by a final state constituted by 6 objects: 4 jets, one charged lepton (electron or muon) and one neutrino. However, at the LHC energies, boosted quarks can be produced. The decay products of boosted

Mass (GeV)	Z' Expected (pb)	Z' Observed (pb)	g_{KK} Expected (pb)	g_{KK} Observed (pb)
Without Systematic Uncertainties				
500	2.60	2.27	3.03	2.50
600	1.93	1.42	2.18	1.63
700	1.43	1.00	1.80	1.21
800	1.11	0.87	1.45	1.14
1000	0.69	1.06	0.86	1.26
1300	0.46	0.58	0.61	0.82
1600	0.33	0.46	0.45	0.70
1800	–	–	0.28	0.34
2000	0.27	0.28	–	–
With Systematic Uncertainties				
500	8.49	9.26	10.34	10.12
600	6.03	4.81	6.00	4.96
700	3.13	2.49	4.21	3.07
800	2.10	1.86	2.68	2.17
1000	1.10	2.35	1.38	2.91
1300	0.62	0.95	0.90	1.64
1600	0.46	0.76	0.68	1.37
1800	–	–	0.41	0.60
2000	0.37	0.40	–	–

Table 6.11: Expected and observed limits on the signal cross-sections [8].

tops can not be resolved individually, since two or more of them can be merged. The selection and reconstruction used in the analysis was designed for the particular resolved topology and events with three jets, to take into account the possibility that one of the jets was not reconstructed in the event or that two of the top-quarks decay products are merged.

The expected contribution of SM processes to the $t\bar{t}$ invariant mass spectra was determined using a combination of Monte Carlo simulated samples and data. Four different methods to reconstruct the top pair were studied: 4 hardest jets, dRmin, χ^2 and χ^2 scaled. It was decided to use the dRmin method as the reconstruction method of the analysis, since the performance of the 4 hardest jets was lower and the χ^2 and χ^2 scaled methods are sensitive to pile-up. The dRmin method removes jets that are "far" from the rest of the activity in the event. The reconstructed $t\bar{t}$ invariant mass is found to be compatible with the SM prediction. 95% C.L. upper limits are set on the production cross-section times branching ratio to top quark pairs of such resonances. Observed limits range from 9.3 pb for a mass of 500 GeV to 0.95 pb for a mass of 1300 GeV, excluding a leptophobic topcolor Z' boson with 500 GeV$< m_{Z'} <$880 GeV. Kaluza-Klein gluons in the Randall Sundrum model with masses between 500 and 1130 GeV are excluded at 95% C.L. Most of the results presented in this chapter are part of the paper in Ref. [4].

Chapter 7

Jet vertex fraction

Contents

7.1 Overview

The number of pile-up interactions increases with increasing luminosity. These additional pp interactions are uncorrelated with the hard process. The presence of pile-up collisions can have basically three effects:

- The measured energy of jets coming from the interesting hard process increases. This jet energy contribution from pile-up is partially corrected through the offset correction described in Section 4.7.4.1.

- Smearing of the jet energy due to the fact that not all jets have the same energy contribution from pile-up. Measurements that are sensitive to the jet energy resolution thus remain dependent on the number of reconstructed primary vertices and/or $< \mu >$ even after the offset correction is applied.

- The presence of additional jets in the hard scatter event coming from additional pp collisions. It makes the hard-scatter jets identification and the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction difficult.

It is of crucial importance to reduce these pile-up effects in order to improve the precision and sensitivity of physics analyses at high luminosities. The Jet Vertex Fraction (JVF) [164] is a variable used in ATLAS that quantifies the fraction of the total track transverse momentum originating from a given primary vertex associated/matched to a given jet. Therefore, JVF can be used to identify the origin vertex of a given jet. A cut in the JVF variable can help to filter jets coming from additional pp collisions in the event. In this chapter the performance of JVF and the measurement of the corresponding scale factors are presented. The data and Monte Carlo samples used in the analysis are described in Section 7.3. It is followed in Section 7.4 by a discussion of the different event selections used along this chapter. The determination of the different backgrounds is described in Section 7.5. The determination of the optimal JVF cut to be used for top quark analyses using data collected in 2011 and its performance in Monte Carlo are presented in Section 7.6. The comparison of the JVF requirement efficiencies/inefficiencies in data and Monte Carlo and the derivation of the corresponding scale factors needed to match the JVF cut efficiencies/inefficiencies in Monte Carlo to the one in data are presented in Section 7.7. In Section 7.8, the different systematic uncertainties affecting the scale factors estimation are listed.

7.2 Jet vertex fraction description

Using the tracks reconstructed based in the ATLAS ID information, the JVF variable can be defined for each jet with respect to each identified PV in the event, by identifying the PV of origin of charged particle tracks pointing towards the given jet . Once the hard-scatter PV is selected the JVF variable can be used to select jets having a high likelihood of originating in that vertex. Tracks were required to have at least 7 SCT+pixel detector hits, at most two holes in the pixel detector, $p_{\mathrm{T}}^{track} > 400$ GeV, $|d_0| < 2.5$ mm (with respect to the PV) and $|z_0 \times \sin(\theta)| < 2$ mm (with respect to the PV). Tracks are matched to calorimeter jets throught an angular matching criteria, i.e. $\Delta R(\text{jet, track}) \leq 0.4$. Then, the JVF is calculated as the ratio of the sum of transverse momentum of matched tracks that originates from a chosen PV to the sum of transverse momentum of all matched tracks in the jet, independently of their origin. JVF is defined for each jet with respect to each PV. For a given jet jet_i, its JVF with respect to the primary vertex PV_j is given by:

$$\mathrm{JVF}(\mathrm{jet}_i, \mathrm{PV}_j) = \frac{\sum_k p_{\mathrm{T}}(\mathrm{track}_k^{\mathrm{jet_i}}, \mathrm{PV}_j)}{\sum_n \sum_l p_{\mathrm{T}}(\mathrm{track}_l^{\mathrm{jet_i}}, \mathrm{PV}_n)}, \tag{7.1}$$

where k runs over all tracks originating from PV_j matched to jet_i, n over all vertices in the event and l over all tracks originating from PV_n matched to jet_i. From now on, JVF will be defined with respect to the event hard-scatter vertex, which is selected as the primary vertex with the highest $\sum_{\mathrm{tracks}}(p_{\mathrm{T}}^2)$. In this way, JVF represents some kind of probability for jets with matched tracks to come from the hard-scatter interaction. The JVF sum over all the PVs can be larger than one, since tracks are allowed to be matched to different PVs, if more than one PV is within $|z_0 \times \sin(\theta)| < 2$ mm. The principle of the JVF variable is shown schematically in Figure 7.1.

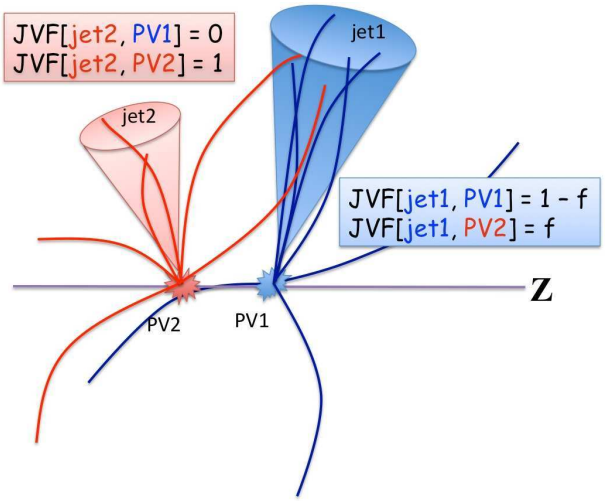


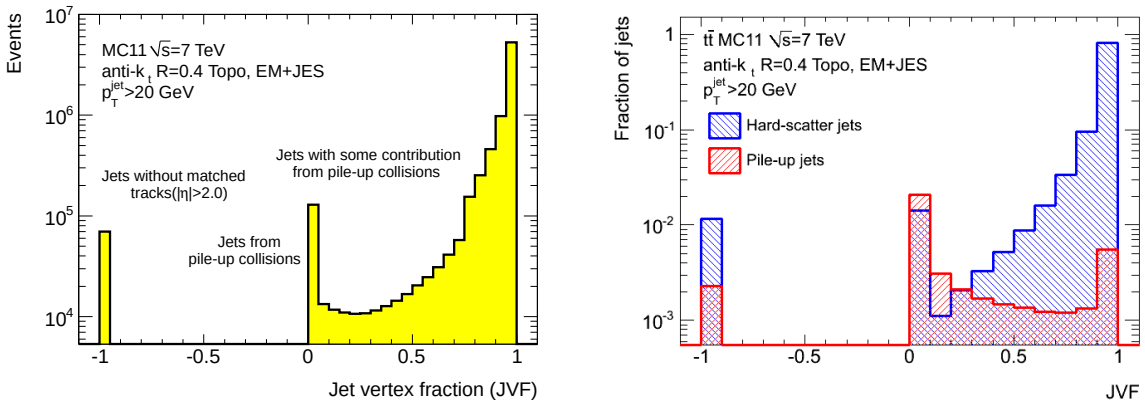
Figure 7.1: Schematic representation of the JVF principle.

The distribution of JVF for jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV in a Monte Carlo $t\bar{t}$ sample in the lepton plus jets channel, which have been reweighted to match the luminosity profile of data collected in 2011, is shown in Figure 7.2a. Three different regions can be distinguished in Figure 7.2a:

- A JVF $= -1$ can be assigned to calorimeter jets which do not have associated tracks. In the studies presented here these jets are accepted, since most of them are jets out of the ID acceptance or neutral-dominated jets which do not have at least 2 matched tracks passing track selection.
- Jets with JVF $= 0$ have all their matched tracks pointing to pile-up vertices. These jets typically have low transverse momentum [164].
- Jets with $0 <$ JVF ≤ 1 have some tracks which originate from a primary interaction other than the selected hard-scattering one. Jets with significant pile-up contribution (JVF closer to 0) will strongly affect measurements in physics analyses. Jets with smaller pile-up contribution (JVF closer to 1) may still exhibit energy scale offsets and angular shifts. Jets with JVF $= 1$ have all of their matched tracks originating from the selected primary hard-scattering vertex. But it does not mean for sure that the jet is completely pile-up contribution free. In an event with a high number of pile-up interactions it is very unlikely that a jet will have no pile-up contributions at all. Therefore, in that case JVF $= 1$ might simply indicate jets with large neutral pile-up composition.

JVF can be used as a tool against pile-up effects in two different ways. The first one is to use a JVF cut to reject pile-up jets. The second one is to use JVF to remove the pile-up contribution to the jet energy measurement on a jet-by-jet basis. As the luminosity increases, the use of a JVF cut will lead to a loss in the jet efficiency selection and to an increase of the jet energy response, as more jets will be affected by pile-up contributions. At that time, the best option will be to keep only those jets that are slighly affected by pile-up and to subtract the contribution to their jet energy from pile-up on a jet-by-jet basis [165]. However, as this kind of techniques was not commisioned in ATLAS at the time of this thesis, many analyses use a JVF cut to reject pile-up jets.

Figure 7.2b shows the JVF distribution for hard-scatter jets and for pile-up jets. It shows the discriminating power of the JVF variable. In Monte Carlo, hard-scatter and pile-up jets are



(a) (b)

Figure 7.2: (a) JVF distribution for jets with a $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV in a Monte Carlo $t\bar{t}$ sample in the lepton plus jets channel. Three different regions can be distinguished in the distribution: jets with some contribution from pile-up ($0 < \mathrm{JVF} \leq 1$), jets from pile-up collisions ($\mathrm{JVF} = 0$), and jets without matched tracks ($\mathrm{JVF} = -1$). (b) JVF distribution for hard-scatter (blue) and pile-up (red) jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV and $|\eta| < 2.5$. Using JVF directly as a discriminating variable provides a way to separate both classes of jets.

defined as follows:

- **Hard-scatter jets:** calorimeter jets matched to truth[1] jets from the hard-scatter ($\Delta R \leq 0.4$).
- **Pile-up jets:** calorimeter jets not matched to truth jets from the hard-scatter ($\Delta R > 0.4$).

The ΔR distribution between the reconstructed jet and the closest truth jet is shown in Figure 7.3a for different JVF ranges. The ratio between the p_{T} of the reconstructed jet and the p_{T} of closest truth jet, $p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{truth}}$, is shown in Figure 7.3b as a function of ΔR. For jets with $0.0 \leq |\mathrm{JVF}| < 0.5$ and at low ΔR, $p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{truth}}$ is bigger than for other JVF ranges as they have extra energy coming from pile-up interactions. Above $\Delta R = 0.4$ the ratio is below one, indicating that most probably the calorimeter jet was wrongly matched.

The performance of a JVF cut can be measured using four different variables:

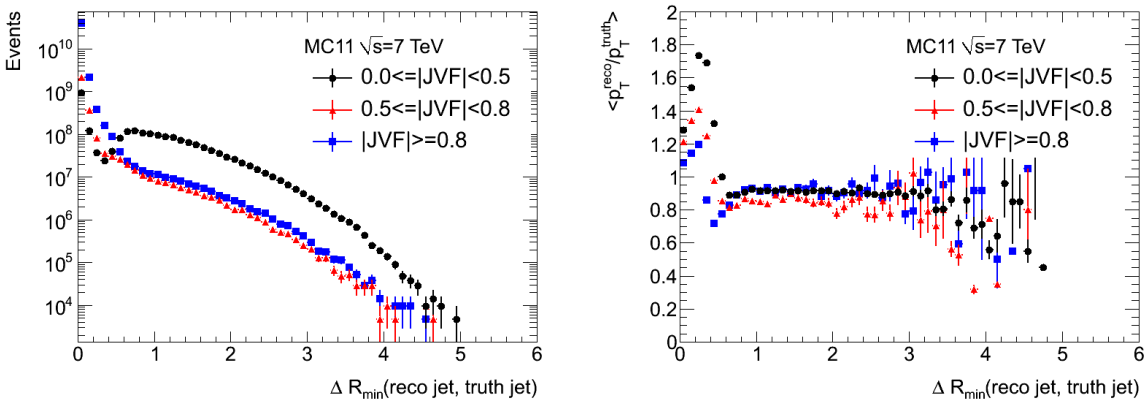
Hard scatter jet selection efficiency $\mathcal{E}_{HS}$: this is the efficiency with which a jet originating from a hard-scatter interaction passes the JVF threshold, being classified as hard-scatter jet.

Pile-up jet rejection $\mathcal{E}_{PU}$: this is the efficiency with which a jet originating from a pile-up interaction fails the JVF selection, being classified as pile-up jet.

Mistag rate for a hard scatter jet $\mathcal{I}_{HS}$: this is the probability of misclassifying a hard-scatter jet as pile-up jet. It is defined as $\mathcal{I}_{HS} = 1 - \mathcal{E}_{HS}$.

Mistag rate for a pile-up jet $\mathcal{I}_{PU}$: this is the probability of misclassifying a pile-up jet as hard-scatter jet. It is defined as $\mathcal{I}_{PU} = 1 - \mathcal{E}_{PU}$.

[1] Truth jets with $p_{\mathrm{T}} > 5$ GeV and $|\eta| < 5$ are considered.



(a) (b)

Figure 7.3: (a) ΔR distribution between the reconstructed jet and the closest truth jet and (b) $p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{truth}}$ as a function of ΔR for different JVF ranges.

The first two are known as JVF efficiency measurements, while the last two are known as JVF inefficiency measurements. Naturally, these efficiencies and inefficiencies are by definition in the $[0, 1]$ range. In order to extract the hard-scatter jet selection efficiency from data a tag & probe method needs to be used. Events where one high-p_{T} Z boson and one back-to-back jet were produced are used to calculate $\mathcal{E}_{HS}$ and $\mathcal{I}_{HS}$. Events where a low-p_{T} Z and exactly only one jet were produced are used to estimate $\mathcal{E}_{PU}$ and $\mathcal{I}_{PU}$. A jet present in an event where a non-boosted Z boson has been produced is frequently coming from a pile-up interaction. The specific event selection will be described in detail in Section 7.4.3 The same tag and probe method can be also used in Monte Carlo. Differences in efficiencies between data and simulation, if any, need to be corrected by using scale factors.

7.3 Data and simulated samples

Two different event selections will be described in this section. The first one corresponds to the one used to select $t\bar{t}$ events in the lepton plus jets channel, from which the optimal JVF cut is derived. The second one is used to select $Z(\rightarrow ee/\mu\mu)$+jets events, from which the scale factors on the JVF requirement are estimated.

7.3.1 Data sample

The data used in this study was recorded by the ATLAS detector between April and October 2011, requiring the presence of at least one muon or electron, depending on the channel under study. The trigger signatures used varied with data taking periods as detailed in Table 7.1. This data sample corresponds to a total integrated luminosity of around 4.7 fb^{-1} [56, 57] of pp collisions at $\sqrt{s} = 7$ TeV after the trigger selection and the selection of runs with important ATLAS subdetectors operational.

7.3.2 Monte Carlo simulation

Monte Carlo simulated samples have been generated at the centre-of-mass energy of 7 TeV. After event generation, all simulated samples were run through the standard GEANT4 [91, 92] simulation of the ATLAS detector and passed throught the same analysis chain as the data.

Data period	$\mathcal{L}$ (pb^{-1})	Electron channel	Muon channel
B-I	1464.6	EF_e20_medium	EF_mu18
J	226.4	EF_e20_medium	EF_mu18_medium
K	590.4	EF_e22_medium	EF_mu18_medium
L-M	2431.7	EF_e22vh_medium1 OR EF_e45_medium1	EF_mu18_medium

Table 7.1: Triggers used in the different data periods in 2011. The number in the trigger name denotes the approximate value of the trigger threshold.

7.3.2.1 Background processes

- The SM $t\bar{t}$ sample was simulated with the MC@NLO generator v4.01 with the CTEQ10 NLO PDF [166] set and a top quark mass of 172.5 GeV. It was showered using Herwig v6.520 in association with Jimmy v4.31. Only events in which at least one of the W bosons decays leptonically were produced, corresponding to a cross-section of 79.01 pb to which a K-factor of 1.146 was applied to account for NNLO corrections from Hathor [150].

- The electroweak single top samples were also simulated using MC@NLO+Herwig/Jimmy but with the CTEQ6.6 PDF set. Leptonic W-boson decays were required for the s- and t-channel processes. For the Wt process, all decays were produced. The cross-sections used are based on approximate NNLO calculations: 64.57 pb (t-channel) [46], 4.63 pb (s-channel) [47] and 15.74 pb (Wt process) [48].

- W and Z+jets samples were generated with the Alpgen v2.14 generator with the CTEQ6L1 PDF set in exclusive bins of parton multiplicity for multiplicities up to four, and inclusively above that. The events were showered with Herwig and Jimmy. Only leptonic vector boson decays were considered ($W \to l\nu_l$, $Z \to l^+l^-$). For Z+jets samples the interference γ^*/Z was taken into account. They were generated with dileptons in the invariant mass range of $40 < m_{\ell\ell} < 2000$ GeV. These samples can contain W/Z+light quark events and W/Z+heavy quark events. Separate W+jets samples with heavy flavour quark production such as $W + b\bar{b}$, $W + c\bar{c}$ and $W + c$, were also produced. They include all posible decays. Double counting was avoided by removing events with b/c-quarks from the W+light jet samples. The cross-sections and K-factors for the different subsamples are given in Tables 7.2 and 7.4. They are normalized to the inclusive NNLO cross section [152]. The normalization of the W+jets yield is derived from data as will be described in Section 7.5.2.

- Diboson samples WW, WZ and ZZ have been generated using Herwig v6.5 with the MRST2007LO* PDF set and Jimmy. A filter is applied requiring the presence of at least one lepton with $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.8$. The cross-sections (K-factors) used for these filtered samples are: 11.50 pb (1.48) for WW production, 3.46 pb (1.60) for WZ production, and 0.97 pb (1.30) for ZZ production. The K-factors are such that the cross-sections agree with the results obtained using the MCFM [87] generator.

Herwig and Jimmy have been tuned to the ATLAS MC11 tune [167]. All samples were simulated including the effects due to in-time and out-of-time pile-up. In addition, events in simulated samples were reweighted so that the distribution of the average number of interactions per bunch crossing matches the one in the data.

7.3.2.2 Signal Processes

Samples for topcolor Z' bosons with different Z' masses were generated with PYTHIA v6.425 using MRST2007LO* PDFs, allowing all three $t\bar{t}$ topologies. These samples were only used to check the impact of the JVF requirement in the Z' invariant mass spectra. The cross-section and K-factors used are not relevant.

Subsample	Cross section (pb)	K-factor
$W \to e\nu$ + 0lp	6930.50	1.196
$W \to e\nu$ + 1lp	1305.30	1.196
$W \to e\nu$ + 2lp	378.13	1.196
$W \to e\nu$ + 3lp	101.86	1.196
$W \to e\nu$ + 4lp	25.68	1.196
$W \to e\nu$ + 5lp	6.99	1.196
$W \to \mu\nu$ + 0lp	6932.40	1.195
$W \to \mu\nu$ + 1lp	1305.90	1.195
$W \to \mu\nu$ + 2lp	378.07	1.195
$W \to \mu\nu$ + 3lp	101.85	1.195
$W \to \mu\nu$ + 4lp	25.72	1.195
$W \to \mu\nu$ + 5lp	7.00	1.195
$W \to \tau\nu$ + 0lp	6932.40	1.195
$W \to \tau\nu$ + 1lp	1304.90	1.195
$W \to \tau\nu$ + 2lp	377.93	1.195
$W \to \tau\nu$ + 3lp	101.96	1.195
$W \to \tau\nu$ + 4lp	25.71	1.195
$W \to \tau\nu$ + 5lp	7.00	1.195

Table 7.2: Cross-sections for the various $W(\to \ell\nu)$+jets subsamples.

Subsample	Cross section (pb)	K-factor
$W + bb$ + 0lp	47.35	1.20
$W + bb$ + 1lp	35.76	1.20
$W + bb$ + 2lp	17.33	1.20
$W + bb$ + 3lp	7.61	1.20
$W + cc$ + 0lp	127.53	1.20
$W + cc$ + 1lp	104.68	1.20
$W + cc$ + 2lp	52.08	1.20
$W + cc$ + 3lp	16.96	1.20
$W + c$ + 0lp	644.4	1.52
$W + c$ + 1lp	205.0	1.52
$W + c$ + 2lp	50.8	1.52
$W + c$ + 3lp	11.4	1.52
$W + c$ + 4lp	2.8	1.52

Table 7.3: Cross-sections for the various $W(\to \ell\nu)$+heavy flavour jets subsamples.

7.4 Event selection

7.4.1 Objects reconstruction and selection

Physics objects are required to satisfy the following criteria:

Electrons of quality Tight++ are used. Electrons are required to be within the acceptance of the electromagnetic calorimeters ($0 < |\eta_{\text{cluster}}| < 2.47$ and excluding $1.37 < |\eta_{\text{cluster}}| < 1.52$) and to have $E_{\text{T}} > 20$ GeV. The cuts in the isolation variables, `EtCone20` and `PtCone30`, as a function of η_{cluster} and E_{T}, are those that garantee that the efficiency for Tight++ electrons in MC simulation is 90%.

Muons are required to be tight combined `Muid` muons within $|\eta| < 2.5$ and with $p_{\text{T}} > 20$ GeV, `etcone20`< 4.0 GeV and `ptcone30`< 2.5 GeV. Muons are required to have $\Delta R > 0.4$ from any jet with $p_{\text{T}} > 25$ GeV and $|\text{JVF}| > 0.75$.

Jets reconstructed from topoclusters with the anti-k_T algorithm with a radius $R = 0.4$ and calibrated using the `EM+JES` scheme are used. The closest jet within $\Delta R < 0.2$ from an

Subsample	Cross-section (pb)	K-factor
$Z \to ee$ + 0lp	668.32	1.25
$Z \to ee$ + 1lp	133	1.25
$Z \to ee$ + 2lp	40.3	1.25
$Z \to ee$ + 3lp	11.2	1.25
$Z \to ee$ + 4lp	2.7	1.25
$Z \to ee$ + 5lp	0.8	1.25
$Z \to \mu\mu$ + 0lp	658	1.25
$Z \to \mu\mu$ + 1lp	133	1.25
$Z \to \mu\mu$ + 2lp	39.6	1.25
$Z \to \mu\mu$ + 3lp	11.1	1.25
$Z \to \mu\mu$ + 4lp	2.8	1.25
$Z \to \mu\mu$ + 5lp	0.8	1.25
$Z \to \tau\tau$ + 0lp	657	1.25
$Z \to \tau\tau$ + 1lp	133	1.25
$Z \to \tau\tau$ + 2lp	40.4	1.25
$Z \to \tau\tau$ + 3lp	11.0	1.25
$Z \to \tau\tau$ + 4lp	2.9	1.25
$Z \to \tau\tau$ + 5lp	0.7	1.25

Table 7.4: Cross-sections for the various $Z(\to \ell\ell)$+jets subsamples.

electron passing the electron selection cuts is removed, since it is likely that they correspond to the same object. Remaining jets are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.5$.

b–jets are selected using the `MV1` algorithm with an operating point of 0.60, which corresponds to a b-tagging efficiency of 70% and a light quark jet rejection factor of 137 in simulated $t\bar{t}$ events.

Missing transverse energy used is the refined $E_{\mathrm{T}}^{\mathrm{miss}}$. The $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement is based on the transverse momenta of all tight++ electrons (to be consistents with the electron definition used in the analysis). See Section 4.9 for more details about the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction.

7.4.2 $t\bar{t}$ events in the lepton plus jets channel

First a set of cuts are applied to clean-up the samples. Then, extra cuts are applied to gradually enhance the $t\bar{t}$ topology in the samples. The event selection is the following:

Clean-up selection

1. Events in data are required to pass the Good Run List selection defined by the ATLAS Top Group. See Section 4.1 for more information about this data quality requirement.

2. The electron or muon trigger (see Table 7.1) had to have fired.

3. At least one primary vertex reconstructed with more than four tracks originating from it is required.

4. Events with noise bursts and data integrity errors in the LAr calorimeter are rejected.

5. If an event contained any jets with $p_{\mathrm{T}} > 20$ GeV which does not fulfill the jet cleaning cuts, the event is rejected. The jet cleaning criteria are the same used in release 16, except that the jet timing was required to be greater than 25 ns. See Section 6.6 for more information.

Signal enhancement selection

1. Events are required to contain exactly one electron with $E_{\mathrm{T}} > 25$ GeV (electron channel) or one muon with $p_{\mathrm{T}} > 25$ GeV (muon channel).

2. The selected lepton must match the online lepton candidate responsible for the trigger decision (see Table 7.1).

3. Events are required to have at least 4 jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV and $|\eta^{\mathrm{jet}}| < 2.5$.

4. In the electron channel, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be larger than 30 GeV and the lepton+$E_{\mathrm{T}}^{\mathrm{miss}}$ transverse mass, M_{T}, larger than 30 GeV. In the muon channel, the requirements are $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}}+M_{\mathrm{T}} > 60$ GeV.

7.4.3 Z+jets events

Clean-up selection

1. The electron or muon trigger (see Table 7.1) had to have fired.

2. Events with noise bursts and data integrity errors in the LAr calorimeter are rejected.

3. At least one primary vertex had to be reconstructed with more than four tracks originating from it. This helps supressing beam-related background contributions and cosmic rays.

4. If an event contained any jets with $p_{\mathrm{T}} > 20$ GeV which does not fulfill the jet cleaning cuts, the event is rejected.

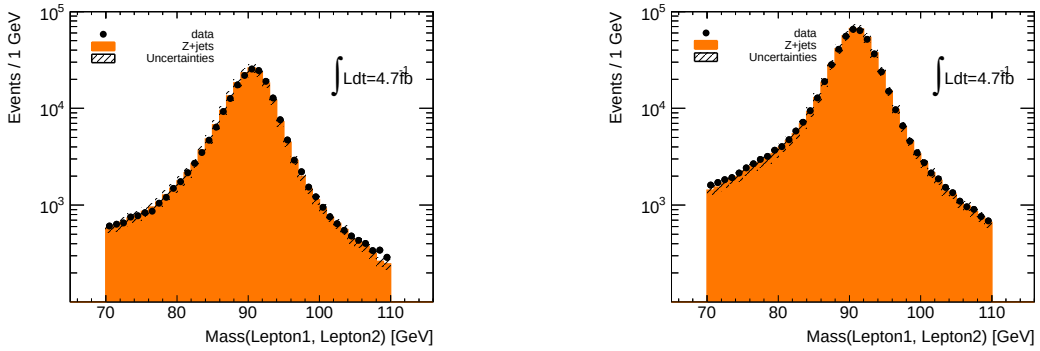
Signal enhancement selection

1. Events are required to contain exactly two electrons with $E_{\mathrm{T}} > 21$ GeV (electron channel) or two muons with $p_{\mathrm{T}} > 20$ GeV (muon channel). No leptons from other flavours are accepted in the event. The E_{T} cut for electrons is set at 21 GeV since the efficiency plateau for the electron trigger efficiency starts at 21 GeV.

2. Both leptons in the event are required to be close to the selected hard-scatter primary vertex. The $|d_0|$ and $|z_0|$ parameters of the corresponding lepton track are required to be smaller than 2 mm.

3. At least one lepton is required to match the trigger lepton (see Table 7.1). The trigger efficiency scale factors applied depend on whether both leptons triggered or only one of the two did.

4. The two leptons are required to have opposite-sign charge and an invariant mass in the range 70 GeV $< M_{\ell\ell} <$ 110 GeV. The Z boson candidate is reconstructed from the four-momenta of the two leptons. The mass distribution of these Z boson candidates in both channels is shown in Figure 7.4 for data and Z+jets events. The contribution from other backgrounds as dibosons, W+jets and $t\bar{t}$ events was checked and found to be negligible.

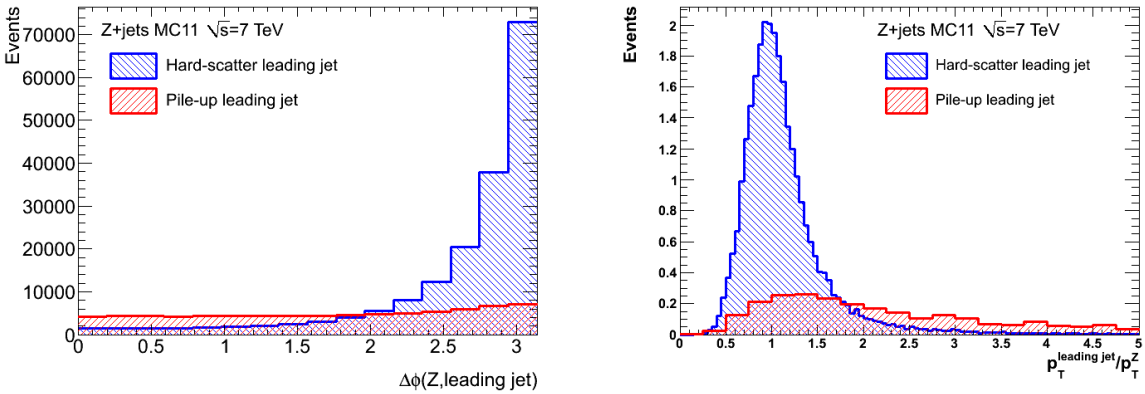


(a) (b)

Figure 7.4: Z boson mass distribution using recontructed (a) $Z \to e^+e^-$ and (b) $Z \to \mu^+\mu^-$ events.

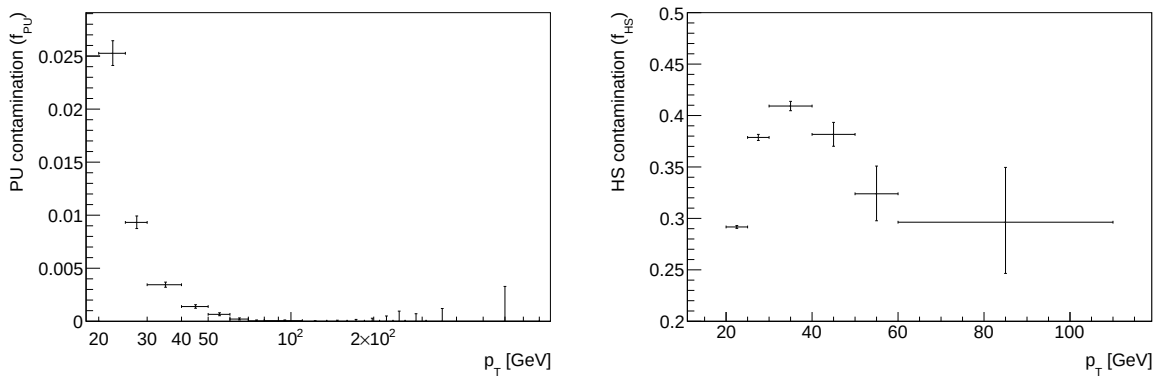
Hard-scatter enriched region The following extra set of cuts aims at selecting events where a boosted Z boson and a jet were produced back-to-back in the hard scatter interaction. This ensures that the jet back-to-back to the Z boson is a hard-scatter jet.

- Events with $p_{\mathrm{T}}^{Z} > 30$ GeV are selected. Only jets with $|\eta| < 2.5$ GeV are considered.
- The highest p_{T} (leading) jet in the event and the Z boson candidates are required to be back-to-back, i.e. $\Delta\phi(Z, \text{leading jet}) > 2.9$. The $\Delta\phi(Z, \text{leading jet})$ distribution before cut is shown in Figure 7.5a when the leading jet is a pile-up or a hard-scatter jet.
- The ratio between the transverse momentum of the leading jet $p_{\mathrm{T}}^{\mathrm{jet}}$ and the transverse momentum of the Z boson candidate p_{T}^{Z} is required to be $0.5 < p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{Z} < 1.5$. This ensures that the jet carries most of the recoil of the Z boson candidate. Figure 7.5b shows the $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{Z}$ distribution for pile-up and hard-scatter leading jets after the applying the cut on the $\Delta\phi(Z, \text{leading jet})$ variable. The pile-up jet contamination fraction is of the order of 2% at low p_{T} and almost zero at high p_{T}, as shown in Figure 7.6a.



(a) (b)

Figure 7.5: (a) $\Delta\phi$ distribution between the Z boson candidate and the leading p_{T} jet in the event and (b) leading p_{T} jet and Z boson candidate p_{T} ratio, when the former was produced in a hard-scatter collision (blue) or in a pile-up collision (red).



(a) (b)

Figure 7.6: (a) Pile-up jet contamination fraction in the hard-scatter enriched region and (b) hard-scatter jet contamination fraction in the pile-up enriched region as a function of the jet p_{T}.

Pile-up enriched region These cuts serve to select events with a Z boson at rest. Any jet in the event will have a high probability of coming from a pile-up interaction.

- Jets with $|\eta| < 4.5$ are selected.
- Events are required to have only one jet with $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV and $|\eta^{\mathrm{jet}}| < 2.5$.
- Events with $p_{\mathrm{T}}^{Z} < 10$ GeV are selected. The p_{T}^{Z} distribution is shown in Figure 7.7 for both cases, when the jet in the event is coming from a pile-up or hard-scatter interaction. Low p_{T}^{Z} region is dominated by pile-up jets. However, a p_{T}-dependent hard-scatter jet contamination fraction is observed in this region, as shown in Figure 7.6b. There is a bump in the p_{T}^{Z} distribution for hard-scatter jets which probably correspond to a mix of pile-up and hard-scatter jets. This feature would need further investigation.

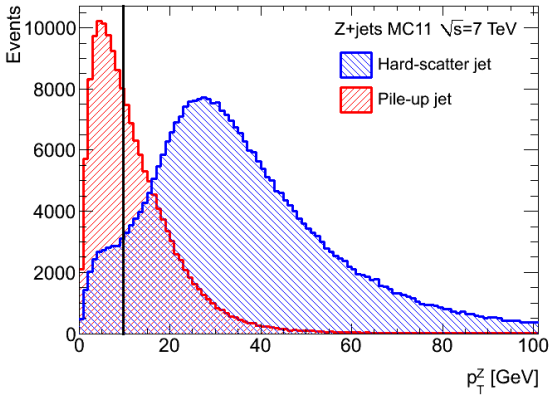


Figure 7.7: Z transverse momentum distribution in events with exactly one jet with $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV and $|\eta^{\mathrm{jet}}| < 2.5$. The distribution is shown when the jet is coming from a hard-scatter (blue) or a pile-up (red) interaction.

7.5 Background estimation: $t\bar{t}$ semileptonic topology

7.5.1 QCD background

The shape of the QCD background is modelled using the jet-electron method [168] (described in more detail in Section 6.7.3). The templates used were produced using the whole data statistics from 2011.

7.5.2 W+jets background

The W+jets background was estimated using the ALPGEN samples described in Section 7.3.2. Same normalization procedure used in Chapter 6 Section 6.7.2. The normalization factors have been updated with respect to the ones used in the analysis described in Chapter 6. The W+jet overall normalization in the tagged 4-jet inclusive bin was found to be 0.83 for the electron channel and 0.82 for the muon channel. And the flavor fraction scale factors for a jet multiplicity of 4 are: 1.09 for the $Wb\bar{b}$ and $Wc\bar{c}$ components, 1.08 for the Wc component and 0.96 for the "light parton" component.

7.6 Optimisation of JVF requirement for top quark analyses

A cut in the JVF variable, cut_{JVF}, can be used to reject pile-up jets in an event. As was already mentioned in Section 7.2, such a cut will have an efficiency for hard-scatter jets and some rejection power for pile-up jets. In order to appropriately select the cut for the $t\bar{t}$ analysis in the lepton plus jets channel, the right compromise between the two needs to be found.

7.6.1 Determination of the optimal JVF requirement

Using the SM $t\bar{t}$ samples described in Section 7.3.2 the discriminating power of JVF is measured by matching calorimeter jets to truth jets. The *hard-scatter jet selection efficiency* is defined as the fraction of hard-scatter jets that satisfy $|\mathrm{JVF}| \geq cut_{\mathrm{JVF}}$. While the *pile-up jet rejection* is defined by the fraction of pile-up jets that have $|\mathrm{JVF}| < cut_{\mathrm{JVF}}$. The absolute value of JVF is used in order to take into account jets with JVF=-1, since these jets can be neutral-dominated jets with less than 2 matched tracks. Figures 7.8a and 7.8b show the inclusive efficiency for retaining hard-scatter jets and the achieved inclusive rejection against jets originating from pile-up interactions as a function of the JVF threshold, respectively. Efficiency and rejection curves depend on the jet topology and kinematic, as well as on the number of interactions in the event as can be seen in Figures 7.9 and 7.10. The optimal JVF threshold is found by maximizing the inclusive hard scatter jet selection efficiency times the inclusive pile-up jet rejection (see Figure 7.8c). The maximum was found for a JVF threshold value of 0.74. Based on these studies, jets are required to have $|\mathrm{JVF}| \geq 0.75$, which gives a 80% of rejection against pile-up jets for a 92% hard-scatter jet selection efficiency for a semileptonic $t\bar{t}$ sample.

7.6.2 JVF requirement performance

Physics analyses must be insensitive to contributions to jet multiplicity from pile-up as these are entirely uncorrelated with the hard-scattering process of interest. Figure 7.11a shows the average jet multiplicity distribution as a function of N_{PV} for jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV and $|\eta| < 2.5$ in the $t\bar{t}$, W+jets and single-top samples when requiring at least 4 jets with $p_{\mathrm{T}} > 25$ GeV. The same plot with the additional requirement of $|\mathrm{JVF}| \geq 0.75$ for all the jets is presented in Figure 7.11b. When requiring $|\mathrm{JVF}| \geq 0.75$ the dependence of the jet multiplicity on N_{PV} is reduced. The

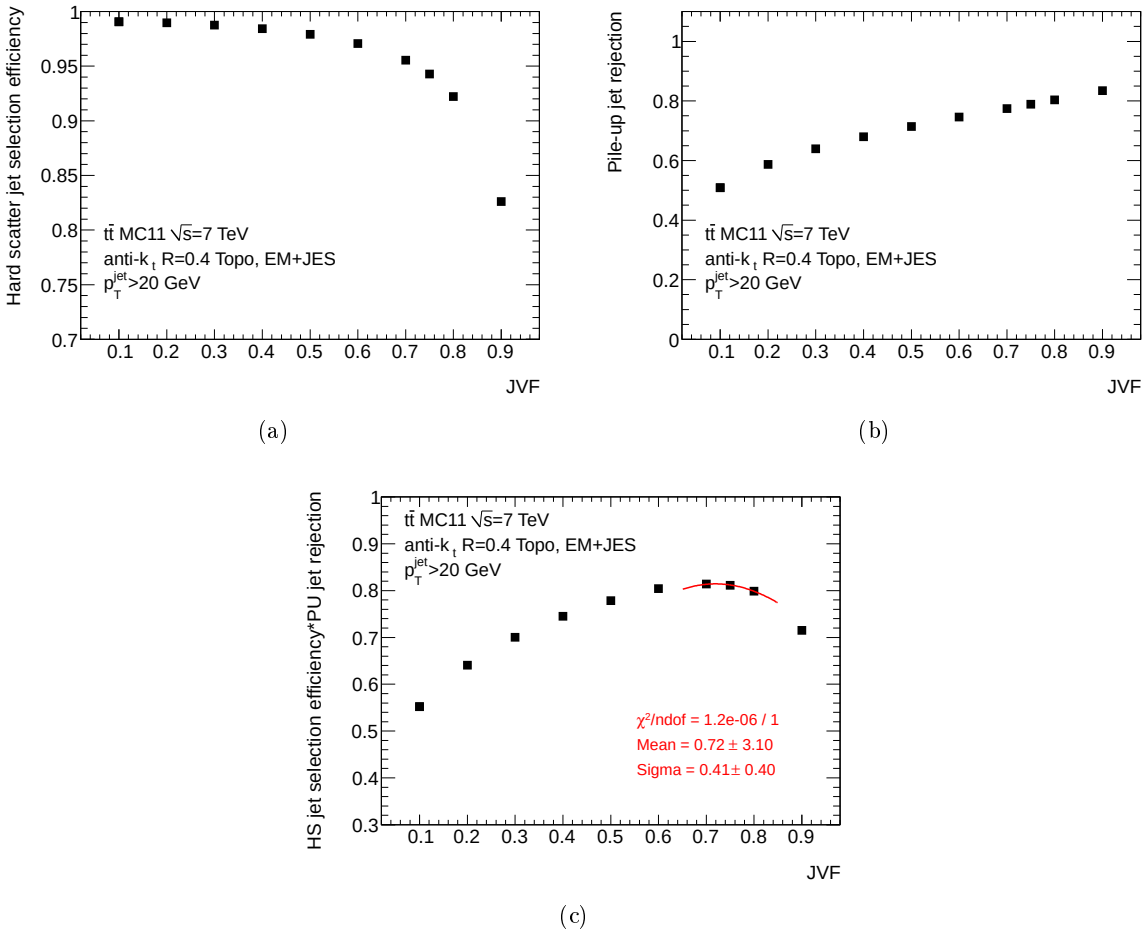


Figure 7.8: (a) Hard-scatter selection, (b) rejection efficiency against pile-up jets and (c) hard-scatter selection efficiency times pile-up rejection efficiency as a function of the |JVF| threshold for jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV and $|\eta| < 2.5$ in the SM $t\bar{t}$ sample. The maximum was estimated with a polynomial fit computed with the 4 points around the maximum sampled point.

decision of the ATLAS top group was to include an additional requirement in the event selection for the $t\bar{t}$ analyses: at least four jets with $p_{\mathrm{T}} > 25$ GeV and $|\mathrm{JVF}| \geq 0.75$.

The impact in the efficiency selection of asking at least 4 jets with $|\mathrm{JVF}| \geq 0.75$ in the $t\bar{t}$ lepton plus jets analysis (same analysis presented in Chapter 6 but using release 17 with 4.7 fb^{-1}, the event selection is described in Section 7.4.2) is shown in Figure 7.12, where the acceptance of the JVF requirement is drawn as a function of the reconstructed $t\bar{t}$ invariant mass for different samples. The $t\bar{t}$ invariant mass was reconstructed using the two methods described in Section 6.9: dRmin and χ^2. Using the JVF requirement allows to reduce the number of W+jets and single-top events especially at low $t\bar{t}$ mass, while keeping an acceptance in the SM $t\bar{t}$ sample bigger than 0.8 in the whole mass spectra. The results indicate that the JVF requirement allows to reduce the main backgrounds for the analysis presented in Chapter 6 especially at low $t\bar{t}$ mass.

In addition, the impact of the event JVF requirement in the $t\bar{t}$ reconstruction was verified. Figure 7.13 and 7.14 shows the reconstructed $t\bar{t}$ invariant mass and corresponding resolution obtained with both algorithms for the SM $t\bar{t}$ sample and a Z' with a mass of 1 TeV. No dramatic differences have been observed before and after applying the JVF selection requirement. A tendency to reconstruct higher masses is observed when requiring only jets with $|\mathrm{JVF}| \geq 0.75$ to

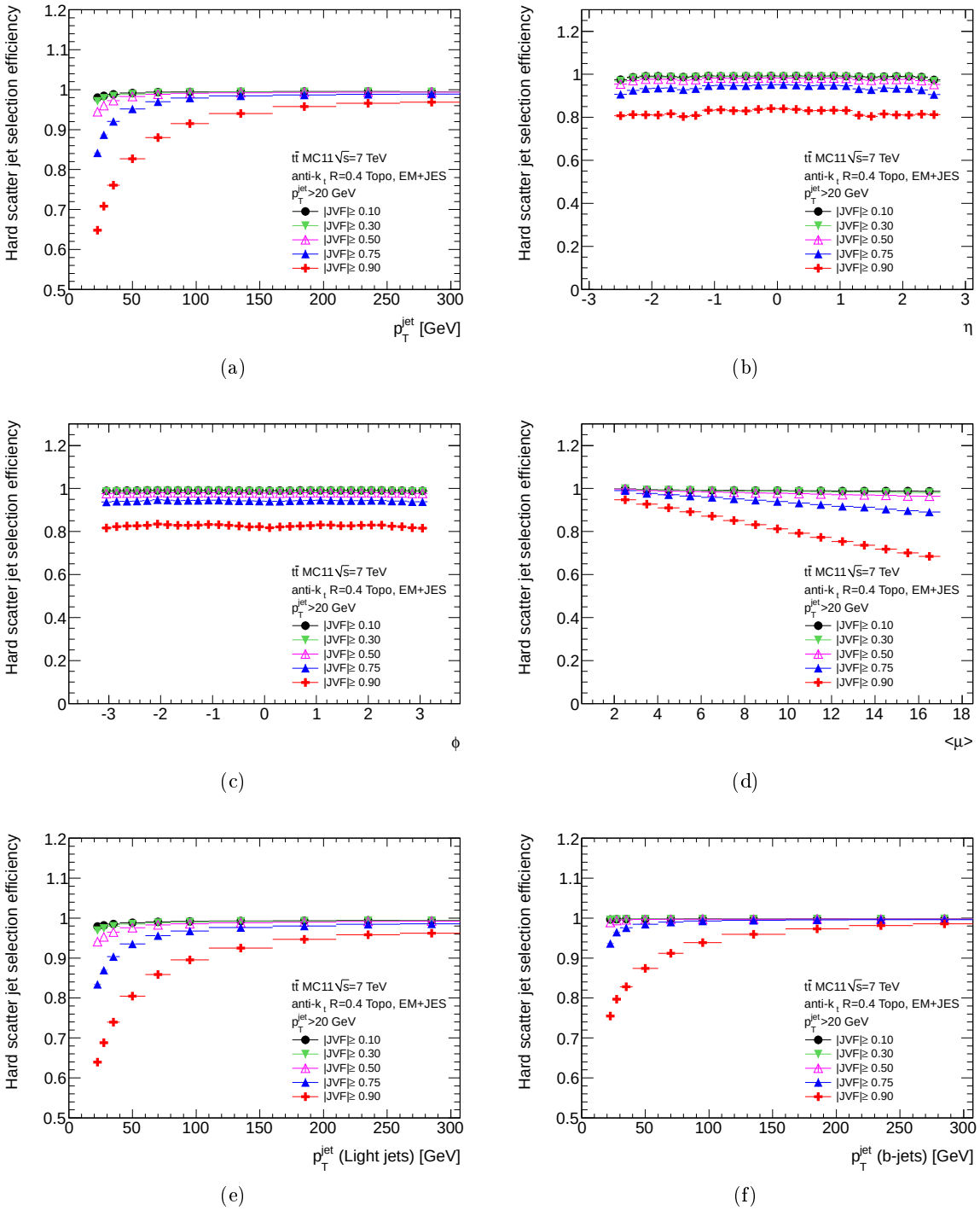


Figure 7.9: Hard-scatter jet selection efficiency as a function of (a) $p_{\mathrm{T}}^{\mathrm{jet}}$, (b) η, (c) ϕ, (d) $< \mu >$ in the inclusive jet sample, (e) $p_{\mathrm{T}}^{\mathrm{jet}}$ for non b-tagged jets and (f) $p_{\mathrm{T}}^{\mathrm{jet}}$ for b-tagged jets. The different curves correspond to different JVF thresholds.

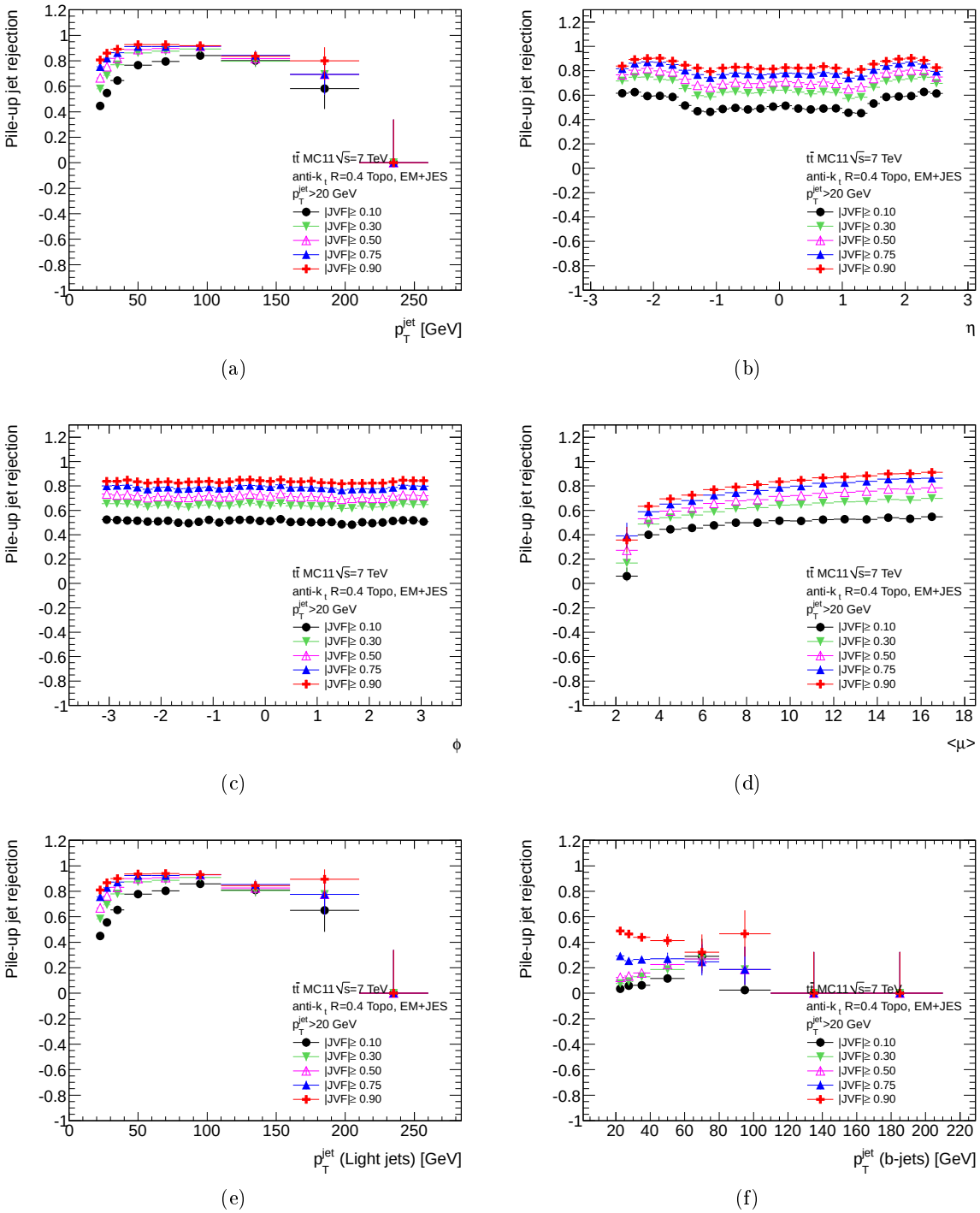


Figure 7.10: Pile-up rejection efficiency as a function of (a) $p_{\mathrm{T}}^{\mathrm{jet}}$, (b) η, (c) ϕ, (d) $< \mu >$ in the inclusive jet sample, (e) $p_{\mathrm{T}}^{\mathrm{jet}}$ for non b-tagged jets and (f) $p_{\mathrm{T}}^{\mathrm{jet}}$ for b-tagged jets. The different curves correspond to different JVF thresholds.

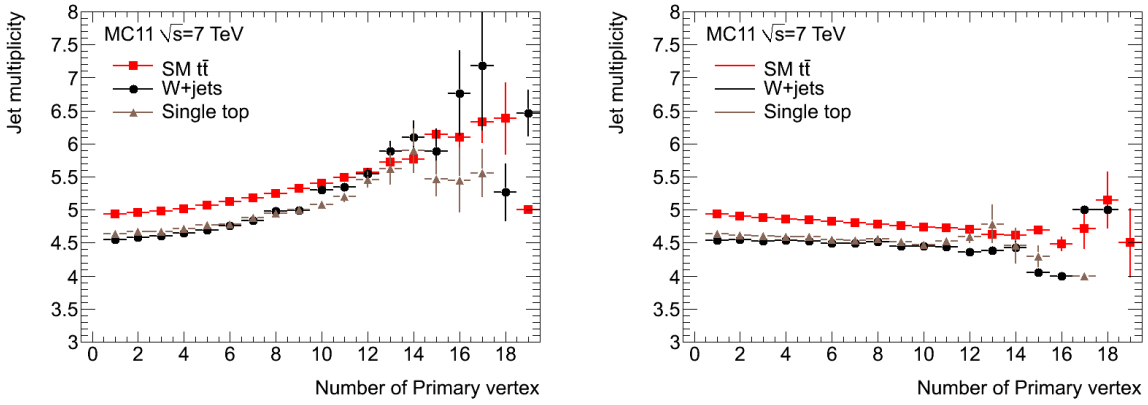


Figure 7.11: Average jet multiplicity as a function of N_{PV} for (a) jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV and $|\eta| < 2.5$ and for (b) jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV, $\eta < 2.5$ and $|\mathrm{JVF}| > 0.75$. At least 4 jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 25$ GeV have been required in both cases.

reconstruct the $t\bar{t}$ pair.

The reconstruction efficiency is shown in Figure 7.15 for the SM $t\bar{t}$ sample, reconstructed using the dRmin and the χ^2 method. The reconstruction efficiency was defined in the same way as in Section 6.9.2. The dependence of the dRmin reconstruction efficiency as a function of $< \mu >$ is small. Its reconstruction efficiency decreases slighly when requiring at least 4 jets with $|\mathrm{JVF}| \geq 0.75$ in the event selection. A similar result is observed for the χ^2 method, but the dependence of the reconstruction efficiency as a function of $< \mu >$ is bigger than the one observed for the dRmin method. This dependence decreases when requiring only using jets with $|\mathrm{JVF}| \geq 0.75$ for the $t\bar{t}$ pair reconstruction.

These results highlight the importance of using a JVF cut requirement in physics analyses to identify and remove the contribution from pile-up interactions and improve the measurement precision in ATLAS analyses.

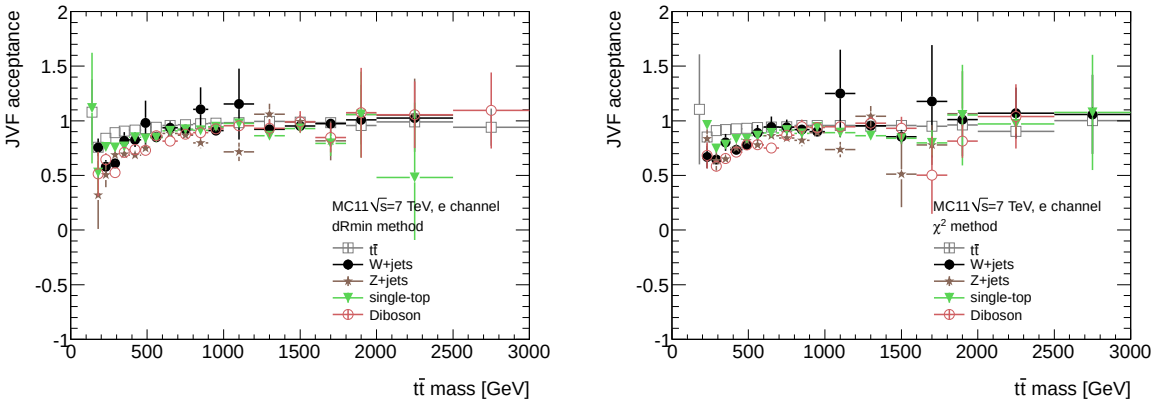
In simulated samples jets can be classified in four different categories:

- $[\mathrm{HS}, \mathrm{JVF}]$ are hard-scatter jets with $|\mathrm{JVF}| \geq 0.75$.
- $[\mathrm{HS}, \overline{\mathrm{JVF}}]$ are hard-scatter jets with $|\mathrm{JVF}| < 0.75$.
- $[\mathrm{PU}, \mathrm{JVF}]$ are pile-up jets with $|\mathrm{JVF}| \geq 0.75$.
- $[\mathrm{PU}, \overline{\mathrm{JVF}}]$ are pile-up jets with $|\mathrm{JVF}| < 0.75$.

In the $t\bar{t}$ semileptonic analysis (see selection described in Section 7.4.2) the fraction of jets (with $p_{\mathrm{T}} > 20$ GeV) in the four different jet categories listed above are given in Table 7.5. W+jets, diboson, Z+jets, $t\bar{t}$ and single top samples were considered. At least four jets with $p_{\mathrm{T}} > 25$ GeV and $|\mathrm{JVF}| \geq 0.75$ were required. The fraction of jets classified as $[\mathrm{HS}, \overline{\mathrm{JVF}}]$, $[\mathrm{PU}, \overline{\mathrm{JVF}}]$ and $[\mathrm{PU}, \mathrm{JVF}]$ is small.

7.7 JVF scale factors

As described in Section 7.2 the performance of JVF can be measured in data using a tag& probe method in Z+jets events. The hard-scatter selection efficiency and the mistag rate for a



(a) (b)

Figure 7.12: JVF requirement acceptance in the $t\bar{t}$ semileptonic analysis as a function of the reconstructed top pair mass. The top pair was reconstructed using two different methods: (a) dRmin and (b) χ^2. The acceptance is shown for the SM $t\bar{t}$, W+jets and single-top samples.

Channel	Electron	Muon
[HS, JVF]	96.1%	96.1%
[HS, $\overline{\text{JVF}}$]	1.9%	2.0%
[PU, $\overline{\text{JVF}}$]	0.9%	0.9%
[PU, JVF]	1.1%	1.0%

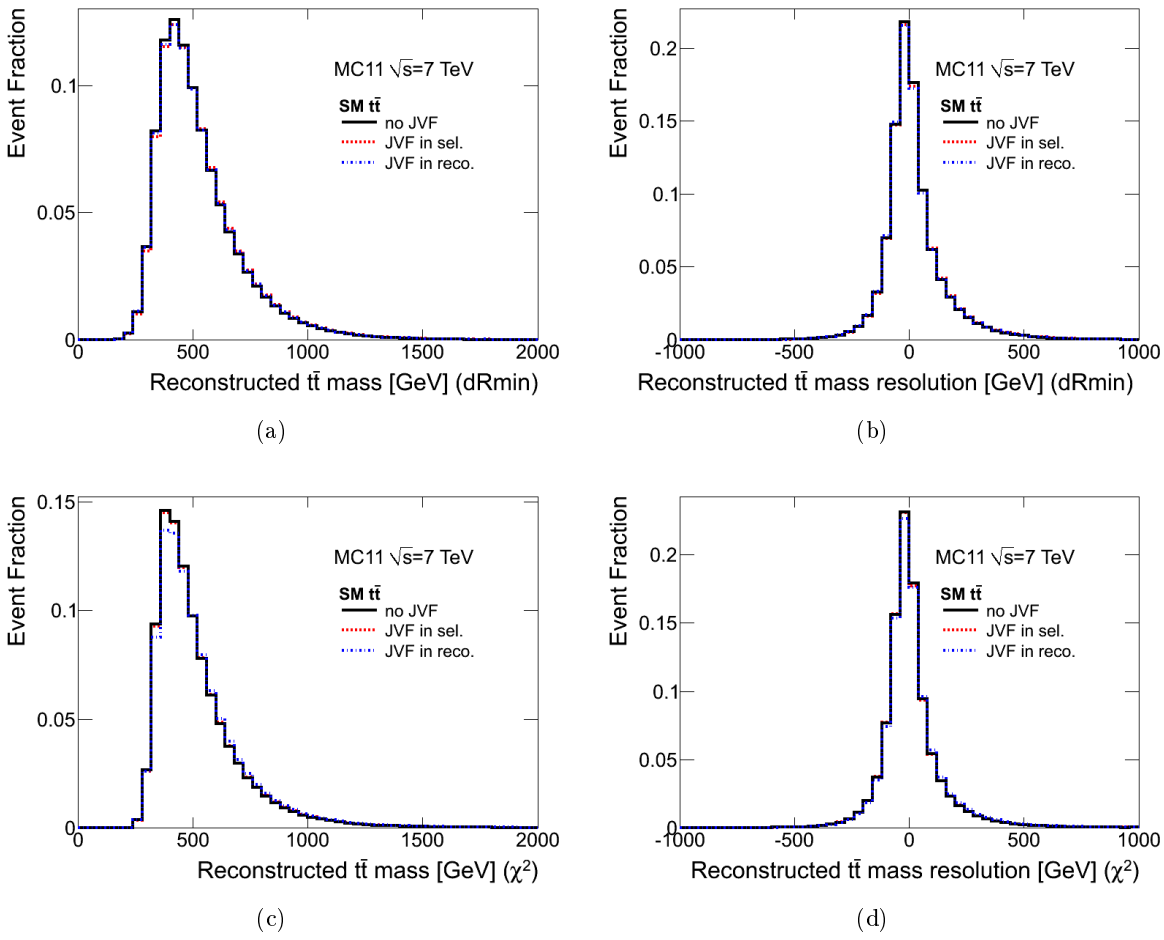
Table 7.5: Fraction of jets (with $p_{\text{T}} > 20$ GeV) classified as [HS, JVF], [HS, $\overline{\text{JVF}}$], [PU, JVF] and [PU, $\overline{\text{JVF}}$] in the $t\bar{t}$ semileptonic analysis.

hard-scatter jet can be measured using events from the hard-scatter enriched region defined in Section 7.4.3. In this scenario, the hard-scatter selection efficiency is defined as the fraction of events where the jet back-to-back to the Z boson candidate passed the JVF threshold of 0.75. In this way $\mathcal{E}_{HS}$ and $\mathcal{I}_{HS}$, defined as $1 - \mathcal{E}_{HS}$, can be calculated from data and Monte Carlo simulation. On the other hand the pile-up rejection efficiency and the mistag rate for a pile-up jet can be determined using events from the pile-up enriched region defined in Section 7.4.3. The pile-up rejection efficiency is defined as the fraction of events where the jet in the event has $|\text{JVF}| < 0.75$. This measured pile-up rejection efficiency suffers from a high hard-scatter jets contamination (as shown in Figure 7.6b) and can be expressed as:

$$\mathcal{E}_{PU}^{\text{measured}} = \frac{\mathcal{E}_{PU} N^{\text{PU}} + \mathcal{I}_{HS} N^{\text{HS}}}{N^{\text{PU}} + N^{\text{HS}}}, \tag{7.2}$$

where:

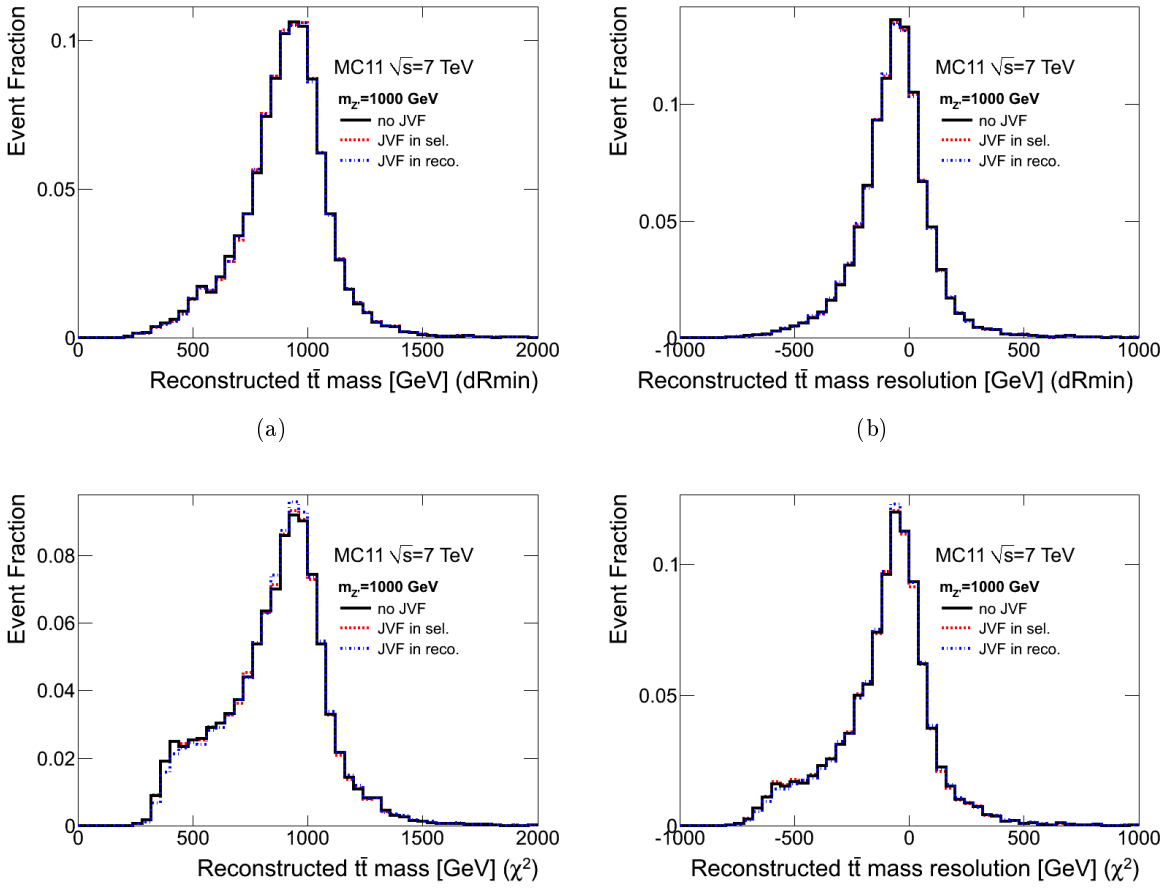
- $\mathcal{E}_{PU}^{\text{measured}}$ is the pile-up rejection efficiency measured in the pile-up enriched region defined in Section 7.4.3.
- $\mathcal{E}_{PU}$ is the real pile-up rejection efficiency. The unknown variable.
- $\mathcal{I}_{HS}$ is the mistag rate for a hard-scatter jet efficiency measured in the hard-scatter enriched region defined in Section 7.4.3.



(a) (b)

(c) (d)

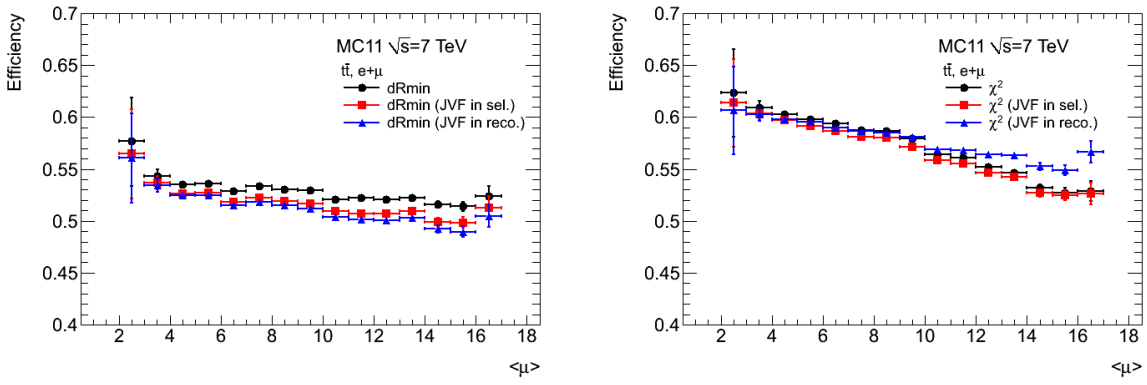
Figure 7.13: Reconstructed $t\bar{t}$ pair invariant mass (left) and its resolution (right) using the dRmin (top) and χ^2 method (bottom) for the SM $t\bar{t}$ when no JVF requirement is applied (black solid line), when requiring at least four jets with $|\text{JVF}| \geq 0.75$ in the event (red dotted line) and when requiring only jets with $|\text{JVF}| \geq 0.75$ to reconstruct the $t\bar{t}$ pair (blue discontinuous line).



(a) (b)

(c) (d)

Figure 7.14: Reconstructed $t\bar{t}$ pair invariant mass (left) and its resolution (right) using the dRmin (top) and χ^2 method (bottom) for the Z' sample with a mass of 1 TeV when no JVF requirement is applied (black solid line), when requiring at least four jets with $|\text{JVF}| \geq 0.75$ in the event (red dotted line) and when requiring only jets with $|\text{JVF}| \geq 0.75$ to reconstruct the $t\bar{t}$ pair (blue discontinuous line).



(a) (b)

Figure 7.15: Reconstruction efficiency using the (a) dRmin and (b) χ^2 method for the SM $t\bar{t}$.

- N^{PU} is the number of events where the jet in the event is classified as a pile-up jet (using the geometric matching criteria defined in Section 7.2).

- N^{HS} is the number of events where the jet in the event is classified as a hard-scatter jet (using the geometric matching criteria defined in Section 7.2).

- $N^{\mathrm{PU}} + N^{\mathrm{HS}}$ is the total number of events in pile-up enriched sample.

Defining f_{HS} as the fraction of hard-scatter jets in the pile-up enriched sample, i.e. the hard-scatter contamination fraction shown in Figure 7.6b:

$$N^{\mathrm{HS}} = f_{\mathrm{HS}}(N^{\mathrm{PU}} + N^{\mathrm{HS}}) \tag{7.3}$$
$$N^{\mathrm{PU}} = (1 - f_{\mathrm{HS}})(N^{\mathrm{PU}} + N^{\mathrm{HS}}). \tag{7.4}$$

$\mathcal{E}_{PU}$ can be expressed as:

$$\mathcal{E}_{PU} = \mathcal{E}_{PU}^{\mathrm{measured}} + \frac{f_{\mathrm{HS}}}{1 - f_{\mathrm{HS}}}(\mathcal{E}_{PU}^{\mathrm{measured}} - \mathcal{I}_{HS}). \tag{7.5}$$

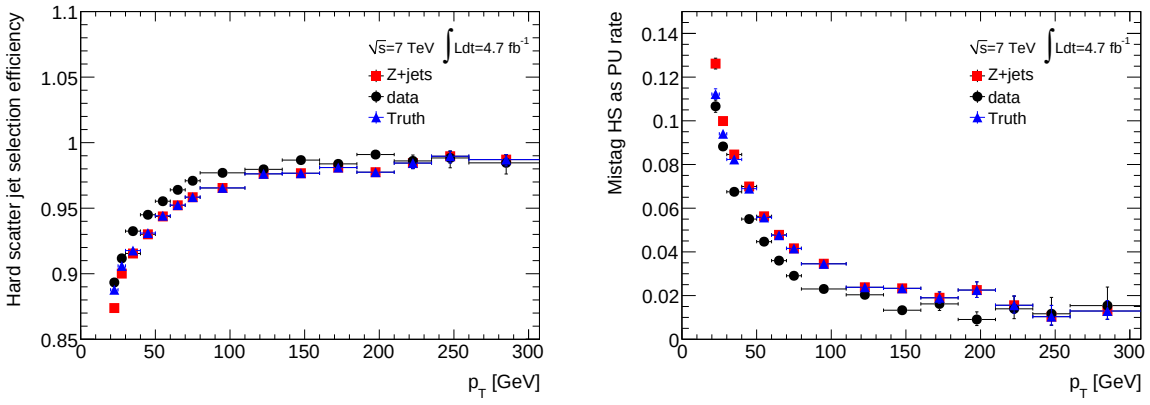
$\mathcal{I}_{PU}$ can be equally calculated since it is defined as $1 - \mathcal{E}_{PU}$. f_{HS} is calculated in simulation and used for both simulation and data.

Figure 7.16 shows $\mathcal{E}_{HS}$ and $\mathcal{I}_{HS}$ in data and Monte Carlo simulation as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ measured using jets with a $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV and $|\eta| < 2.5$. The truth hard scatter jet selection efficiency, $\mathcal{E}_{HS}^{\mathrm{truth}}$, and the truth mistag rate for a hard scatter jet, $\mathcal{I}_{HS}^{\mathrm{truth}}$, calculated in the Z+jets simulated samples are also shown for reference. The JVF variable is not well modelled in the simulation. The Monte Carlo predicts more jets with a high contribution from pile-up than what is seen in data. This causes a difference in the JVF requirement performance in data and simulation. $\mathcal{E}_{HS}^{\mathrm{truth}}$ is defined as the fraction of events where the jet back-to-back to the Z boson candidate classified as hard-scatter jet passed the JVF threshold of 0.75. While, $\mathcal{I}_{HS}^{\mathrm{truth}}$ corresponds to the fraction of events where the jet back-to-back to the Z boson candidate classified as hard-scatter jet has $|\mathrm{JVF}| < 0.75$. The agreement between $\mathcal{E}_{HS}^{\mathrm{truth}}$ and $\mathcal{E}_{HS}$ and between $\mathcal{I}_{HS}^{\mathrm{truth}}$ and $\mathcal{I}_{HS}$ in the Z+jets sample is good, except in the first two bins due to the small pile-up contamination fraction in that region (see Figure 7.6a).

Figure 7.17 shows $\mathcal{E}_{PU}$ and $\mathcal{I}_{PU}$ in data and Monte Carlo simulation as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$. The truth pile-up jet rejection efficiency, $\mathcal{E}_{PU}^{\mathrm{truth}}$, and the truth mistag rate for a pile-up jet, $\mathcal{I}_{PU}^{\mathrm{truth}}$, calculated in the Z+jets simulated samples are also shown for reference. $\mathcal{E}_{PU}^{\mathrm{truth}}$ is defined as the fraction of events where the jet in the event classified as pile-up jet has $|\mathrm{JVF}| < 0.75$. On the other hand, $\mathcal{I}_{PU}^{\mathrm{truth}}$ is defined as the fraction of events where the jet in the event classified as pile-up jet has $|\mathrm{JVF}| \geq 0.75$. The difference between the truth and the tag & probe pile-up rejection efficiency and inefficiency seems to indicate that $\mathcal{I}_{HS}$ calculated in the hard-scatter enriched sample is smaller than the actual $\mathcal{I}_{HS}$ in the pile-up enriched region.

The performance of the JVF requirement in Monte Carlo simulation needs to be calibrated to data. This is done in the form of scale factors, defined as the ratio of the JVF efficiencies or inefficiencies in data and simulation:

$$\kappa_{\mathcal{E}_{HS}} = \frac{\mathcal{E}_{HS}^{\mathrm{data}}}{\mathcal{E}_{HS}^{\mathrm{MC}}} \qquad \kappa_{\mathcal{E}_{PU}} = \frac{\mathcal{E}_{PU}^{\mathrm{data}}}{\mathcal{E}_{PU}^{\mathrm{MC}}} \qquad \kappa_{\mathcal{I}_{HS}} = \frac{\mathcal{I}_{HS}^{\mathrm{data}}}{\mathcal{I}_{HS}^{\mathrm{MC}}} \qquad \kappa_{\mathcal{I}_{PU}} = \frac{\mathcal{I}_{PU}^{\mathrm{data}}}{\mathcal{I}_{PU}^{\mathrm{MC}}}. \tag{7.6}$$



(a) (b)

Figure 7.16: Hard-scatter (a) efficiency and (b) inefficiency in data (black) and Monte Carlo simulation (red) as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$. The truth hard-scatter efficiency and inefficiency (blue) are also shown for reference. Uncertainties are statistical only. Both samples $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$+ jets were combined.

7.7.1 Hard scatter selection efficiency/inefficiency scale factors

The data-to-simulation scale factors, $\kappa_{\mathcal{E}_{HS}}$ and $\kappa_{\mathcal{I}_{HS}}$, as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ are shown in Figure 7.18. Scale factors calculated using $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ events were found to be compatible. They have been parametrized in the following way:

$$\kappa_{\mathcal{E}_{HS}} = a * e^{b \times p_{\mathrm{T}}^{\mathrm{jet}}} + c \tag{7.7}$$

$$\kappa_{\mathcal{I}_{HS}} = a * e^{b \times p_{\mathrm{T}}^{\mathrm{jet}}} + c. \tag{7.8}$$

The fit curve is also shown in Figure 7.18. A small discrepancy between data and simulation as a function of the jet η was also found as shown in Figure 7.19. However, the statistics was not enough to derive bidimensional scale factors. It remains to be done in future studies.

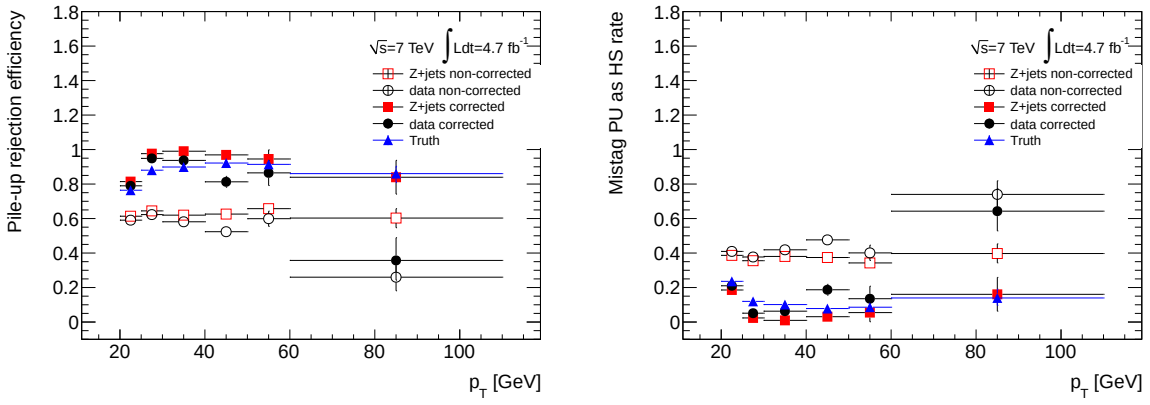
7.7.2 Pile-up rejection efficiency/inefficiency scale factors

The pile-up rejection efficiency scale factors, $\kappa_{\mathcal{E}_{PU}}$, are shown in Figure 7.20. Scale factors calculated using $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ events were found to be compatible. They are compatible with 1, in particular at low p_{T}, but the available statistics is low. The pile-up rejection inefficiency scale factors, $\kappa_{\mathcal{I}_{PU}}$, are not shown here since they reach very high values as $\mathcal{I}_{PU}$ fluctuate very close to 0 (see Figure 7.17).

Given the facts that:

- $\kappa_{\mathcal{E}_{PU}}$ are close to 1.
- It is difficult to derive reliable scale factors for pile-up jets as the statistic is low and the agreement between truth pile-up efficiency/inefficiency and the derived one is not good.
- $\kappa_{\mathcal{E}_{PU}}$ and $\kappa_{\mathcal{I}_{PU}}$ affect only a small fraction of jets in the standard $t\bar{t}$ semileptonic analysis since the fraction of jets classified as [PU, JVF] and [PU, $\overline{\mathrm{JVF}}$] is low (see Table 7.5).

it was decided to set $\kappa_{\mathcal{E}_{PU}}$ and $\kappa_{\mathcal{I}_{PU}}$ to 1.



(a) (b)

Figure 7.17: Pile-up rejection (a) efficiency and (b) inefficiency in data (black) and Monte Carlo simulation (red) as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$. The truth pile-up efficiency and inefficiency (blue) are also shown for reference. The curves before and after correction are shown. Uncertainties are statistical only. Both samples $Z \to ee$ and $Z \to \mu\mu+$ jets were combined.

7.7.3 Impact of the JVF scale factors in $t\bar{t}$ events

Using the scale factors calculated in the previous sections, an event weight can be computed as the product of all the jet scale factors for the jets under consideration. The event weight is given by:

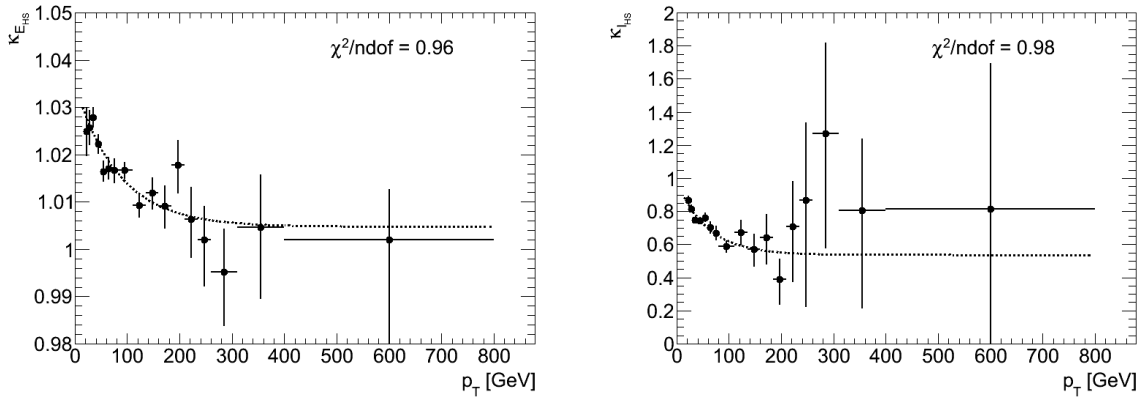
$$w_{\mathrm{event}} = \prod_{N_{\mathrm{HS,JVF}}} \kappa_{\mathcal{E}_{HS}} \prod_{N_{\mathrm{HS},\overline{\mathrm{JVF}}}} \kappa_{\mathcal{I}_{HS}} \prod_{N_{\mathrm{PU},\overline{\mathrm{JVF}}}} \kappa_{\mathcal{E}_{PU}} \prod_{N_{\mathrm{PU,JVF}}} \kappa_{\mathcal{I}_{PU}}, \tag{7.9}$$

where $N_{\mathrm{HS,JVF}}$ is the number of hard-scatter jets with $|\mathrm{JVF}| \geq 0.75$, $N_{\mathrm{HS},\overline{\mathrm{JVF}}}$ is the number of hard-scatter jets with $|\mathrm{JVF}| < 0.75$, $N_{\mathrm{PU,JVF}}$ is the number of pile-up jets with $|\mathrm{JVF}| \geq 0.75$ and $N_{\mathrm{PU},\overline{\mathrm{JVF}}}$ is the number of pile-up jets with $|\mathrm{JVF}| < 0.75$. This event weight has been used in the $t\bar{t}$ semileptonic analysis. Figure 7.21 shows the agreement between data and simulation in the $t\bar{t}$ lepton plus jets analysis before applying any JVF cut, i.e. up to cut 9 of the list presented in Section 7.4.2. The same kinematic distributions are shown after applying the JVF requirement, i.e. at least four jets with $p_{\mathrm{T}} > 25$ GeV and $|\mathrm{JVF}| \geq 0.75$, in Figure 7.22. The disagreement observed between data and simulation is generated by the difference in the JVF performance for data and simulation. Figure 7.23 shows the same kinematic distributions once the JVF event weight has been applied. The agreement between data and Monte Carlo is recovered after applying the JVF event weight.

7.8 Evaluation of the systematic uncertainty associated to the JVF scale factors

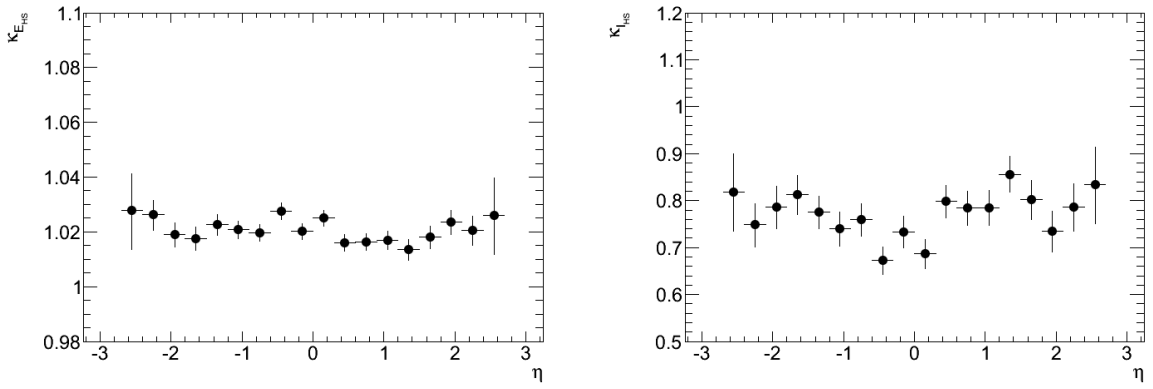
Two different sources of systematic uncertainties have been studied to evaluate the systematic uncertainty associated to $\kappa_{\mathcal{E}_{HS}}$ and $\kappa_{\mathcal{I}_{HS}}$. The systematic uncertainty associated to $\kappa_{\mathcal{E}_{PU}}$ and $\kappa_{\mathcal{I}_{PU}}$ correspond to twice the one calculated for $\kappa_{\mathcal{E}_{HS}}$ and $\kappa_{\mathcal{I}_{HS}}$, respectively, in order to be conservative.

Fit uncertainty: In order to evaluate a systematic uncertainty for the quality of the scale



(a) (b)

Figure 7.18: Hard-scatter (a) efficiency and (b) inefficiency scale factors as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$.



(a) (b)

Figure 7.19: Hard-scatter (a) efficiency and (b) inefficiency scale factors as a function of η^{jet}.

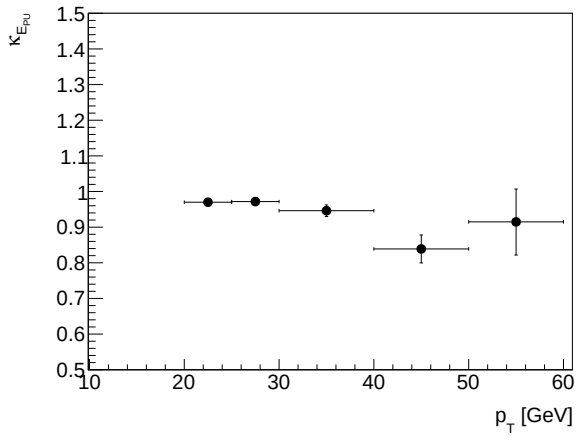


Figure 7.20: Pile-up efficiency scale factors as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$.

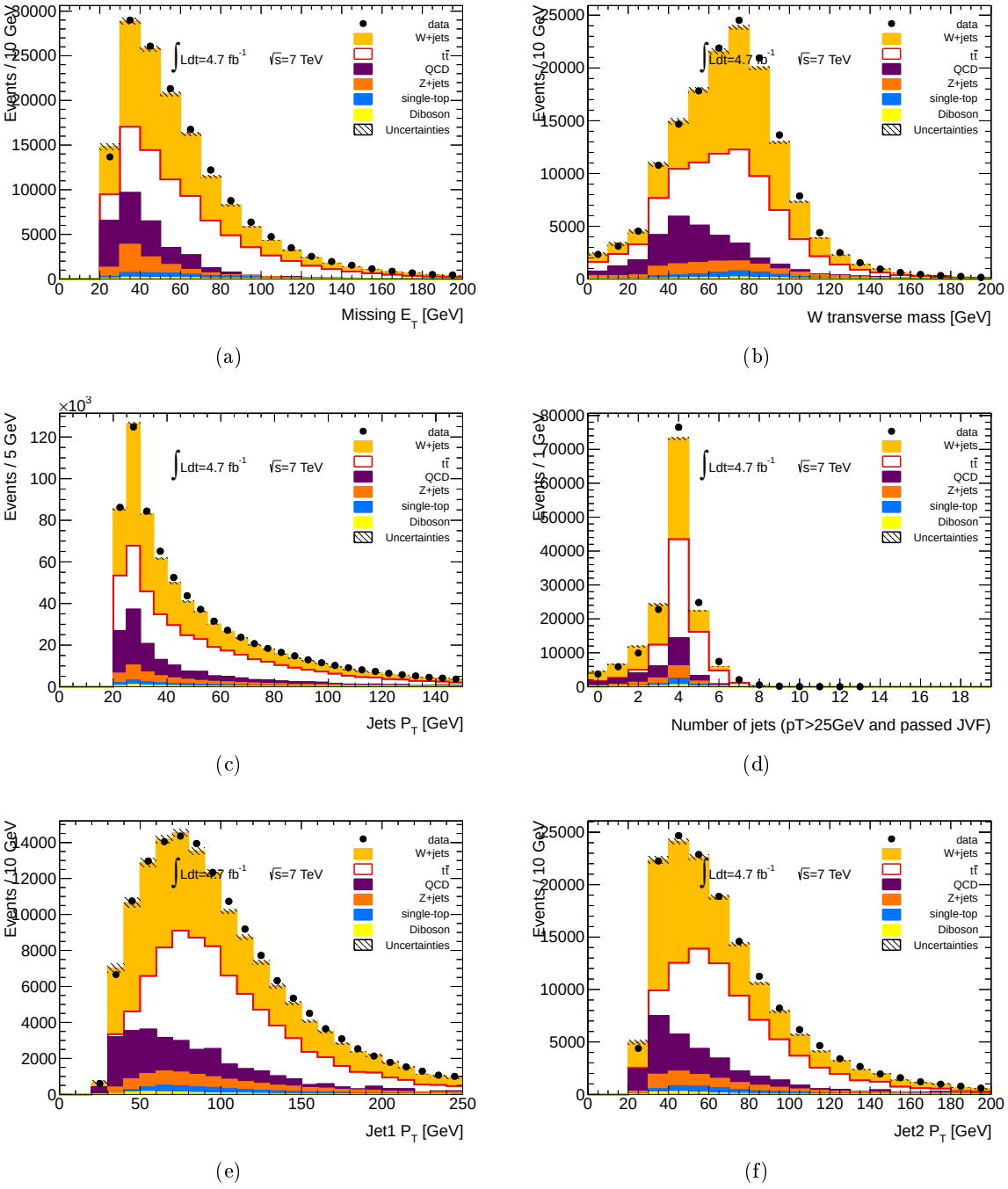


Figure 7.21: (a) $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) lepton-$E_{\mathrm{T}}^{\mathrm{miss}}$ transverse mass, (c) jets p_{T}, (d) number of jets with $p_{\mathrm{T}} > 20$ GeV and $|\mathrm{JVF}| \geq 0.75$, (e) leading jet p_{T} and (f) second jet p_{T} before applying the JVF requirement. The grey hashed area shows the total normalization uncertainty.

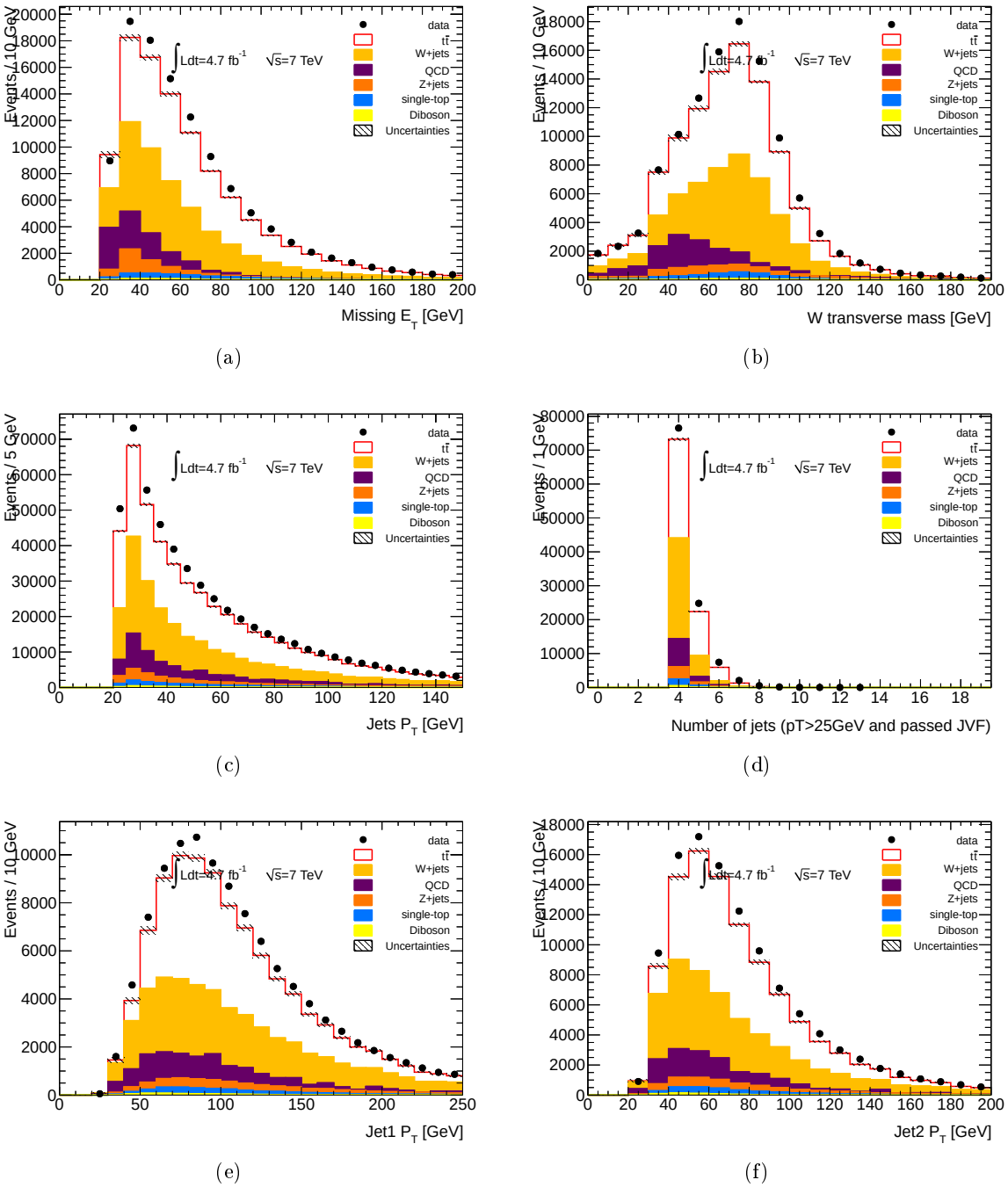


Figure 7.22: (a) $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) lepton-$E_{\mathrm{T}}^{\mathrm{miss}}$ transverse mass, (c) jets p_{T}, (d) number of jets with $p_{\mathrm{T}} > 20$ GeV and $|\mathrm{JVF}| \geq 0.75$, (e) leading jet p_{T} and (f) second jet p_{T} after requiring at least 4 jets with $|\mathrm{JVF}| \geq 0.75$ in the event. The grey hashed area shows the total normalization uncertainty.

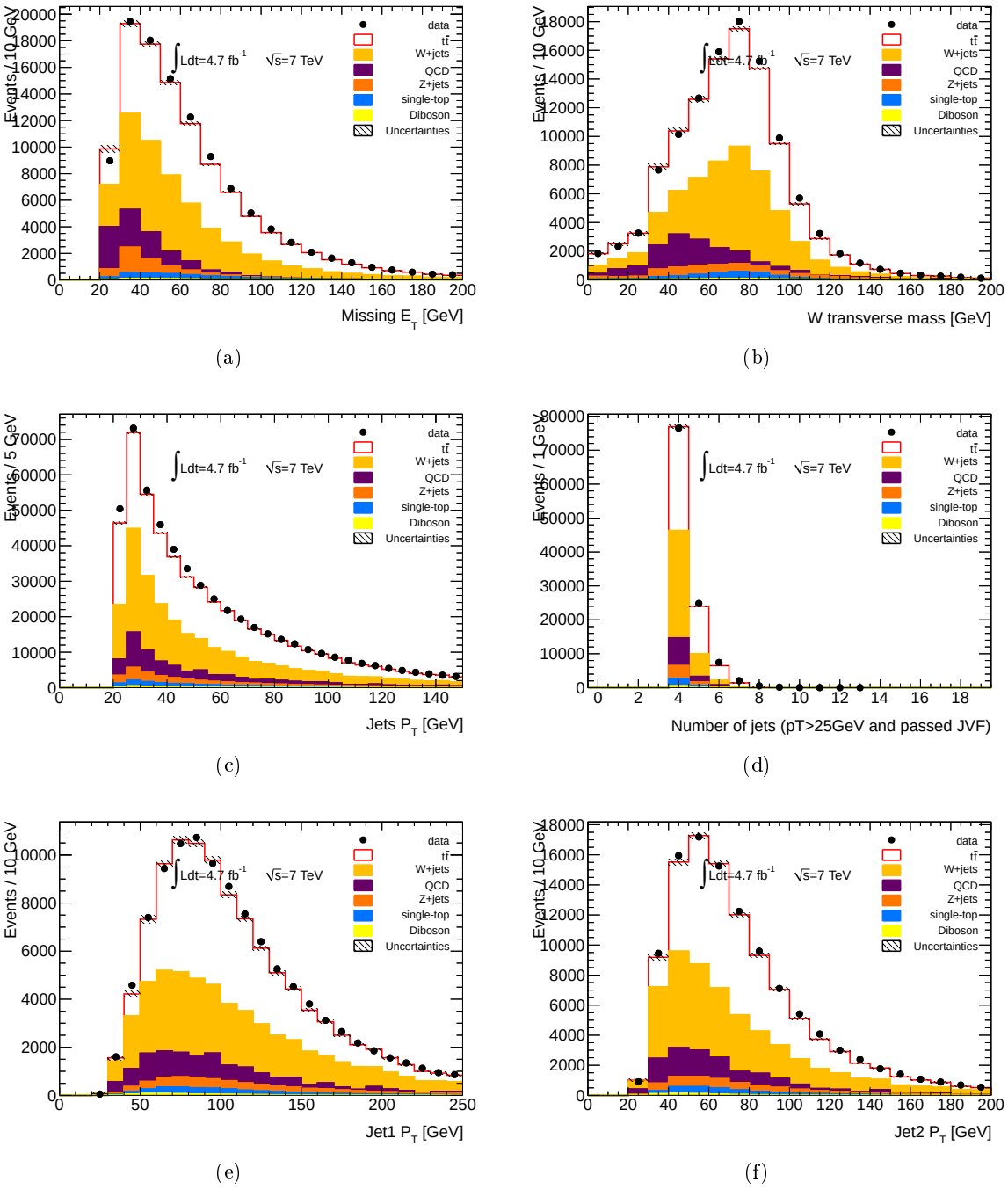
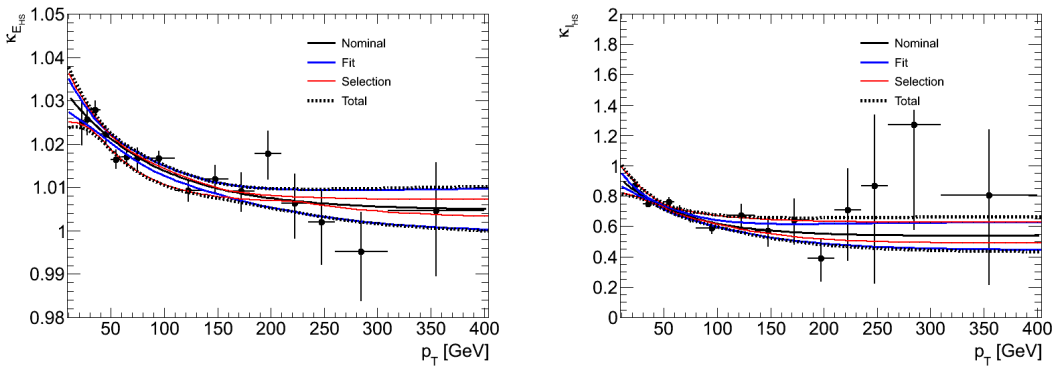


Figure 7.23: (a) $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) lepton-$E_{\mathrm{T}}^{\mathrm{miss}}$ transverse mass, (c) jets p_{T}, (d) number of jets with $p_{\mathrm{T}} > 20$ GeV and $|\mathrm{JVF}| \geq 0.75$, (e) leading jet p_{T} and (f) second jet p_{T} after requiring at least 4 jets with $|\mathrm{JVF}| \geq 0.75$ in the event and applying the JVF scale factors. The grey hashed area shows the total normalization uncertainty.



(a) (b)

Figure 7.24: Systematic uncertainties on (a) $\kappa_{\mathcal{E}_{HS}}$ and (b) $\kappa_{\mathcal{I}_{HS}}$.

factors fits, the uncertainty on a given fit result is enlarged according to the residuals of the fit. The fit uncertainty is calculated as:

$$\varepsilon_{fit} = \max\left\{\sqrt{\frac{\chi^2}{n_d}}, 1\right\} \times \sum_i \sum_j \frac{\partial \kappa_{SF}}{\partial x_i} V_{ij} \frac{\partial \kappa_{SF}}{\partial x_j}, \tag{7.10}$$

where x corresponds to one of the parameters used in the function fit of each scale factor, κ_{SF}, V_{ij} is the covariance matrix between the parameters, n_d is the number of degrees of freedom of the fit, χ^2 is the residual of the fit, and i, j go from 1 to the number of parameters used in the function fit.

Selection uncertainty: The selection criteria used to define the hard-scatter sample (see Section 7.4.3) were varied to check the impact of the selection in the JVF scale factors. For $\kappa_{\mathcal{E}_{HS}}$ and $\kappa_{\mathcal{I}_{HS}}$ two selection criteria were varied:

- $\Delta\phi$ between the leading p_{T} jet and the Z boson candidate is required to be greater than 2.7 and 3.1. The nominal value used was 2.9.
- The cut on p_{T}^{Z} is varied to 20 GeV and 40 GeV. The nominal value corresponds to 30 GeV.

The resulting scale factors, i.e. after variations, are fitted again. For $\kappa_{\mathcal{E}_{HS}}$ the absolute value of the difference between the up variation fit and the nominal fit are taken as the up selection uncertainty. The same holds to calculate the down selection uncertainty. For $\kappa_{\mathcal{I}_{HS}}$ the up and down variations are defined in the other way around in order to take into account the anticorrelation between $\kappa_{\mathcal{E}_{HS}}$ and $\kappa_{\mathcal{I}_{HS}}$.

The fit and selection uncertainty are independent. Therefore, they are added in quadrature to obtain the final systematic uncertainty for a given JVF scale factor. Figure 7.24 summarizes the different contributions to the total uncertainty for $\kappa_{\mathcal{E}_{HS}}$ and $\kappa_{\mathcal{I}_{HS}}$. The impact of the JVF scale factors systematic uncertainty in the SM $t\bar{t}$ samples is shown in Figure 7.25. Variations of the order of 2.5-3% are observed. The event weight corresponding to the up and down systematic uncertainty variations are given by:

$$w_{\text{up/down}} = w_{\text{nom}} \left(1 \pm \sqrt{ \left(\frac{\Delta\kappa^{\text{up/down}}_{\mathcal{E}_{HS},\mathcal{I}_{HS}}}{\prod_{N_{\text{HS,JVF}}} \kappa^{\text{nom}}_{\mathcal{E}_{HS}} \prod_{N_{\text{HS},\overline{\text{JVF}}}} \kappa^{\text{nom}}_{\mathcal{I}_{HS}}} \right)^2 + \left(\frac{\Delta\kappa^{\text{up/down}}_{\mathcal{E}_{PU},\mathcal{I}_{PU}}}{\prod_{N_{\text{PU},\overline{\text{JVF}}}} \kappa^{\text{nom}}_{\mathcal{E}_{PU}} \prod_{N_{\text{PU,JVF}}} \kappa^{\text{nom}}_{\mathcal{I}_{PU}}} \right)^2 } \right), \tag{7.11}$$

where w_{nom} is the nominal event weight and:

$$\Delta\kappa^{\text{up/down}}_{\mathcal{E}_{HS},\mathcal{I}_{HS}} = \prod_{N_{\text{HS,JVF}}} \kappa^{\text{up/down}}_{\mathcal{E}_{HS}} \times \prod_{N_{\text{HS},\overline{\text{JVF}}}} \kappa^{\text{up/down}}_{\mathcal{I}_{HS}} - \prod_{N_{\text{HS,JVF}}} \kappa^{\text{nom}}_{\mathcal{E}_{HS}} \times \prod_{N_{\text{HS},\overline{\text{JVF}}}} \kappa^{\text{nom}}_{\mathcal{I}_{HS}} \tag{7.12}$$

$$\Delta\kappa^{\text{up/down}}_{\mathcal{E}_{PU},\mathcal{I}_{PU}} = \prod_{N_{\text{PU},\overline{\text{JVF}}}} \kappa^{\text{up/down}}_{\mathcal{E}_{PU}} \times \prod_{N_{\text{PU,JVF}}} \kappa^{\text{up/down}}_{\mathcal{I}_{PU}} - \prod_{N_{\text{PU,JVF}}} \kappa^{\text{nom}}_{\mathcal{E}_{PU}} \times \prod_{N_{\text{PU,JVF}}} \kappa^{\text{nom}}_{\mathcal{I}_{PU}}. \tag{7.13}$$

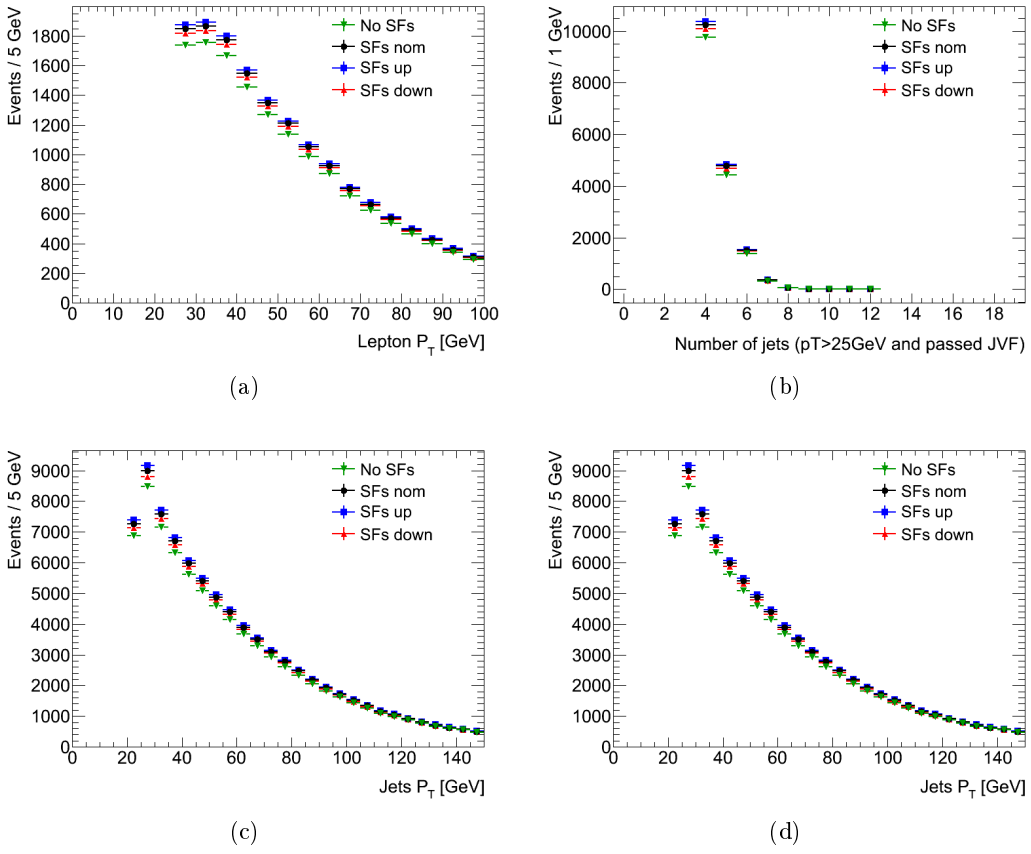


Figure 7.25: Impact of the JVF scale factors systematic uncertainty on the (a) lepton p_{T}, (b) number of jets with $p_{\text{T}} > 20$ GeV and $|\text{JVF}| \geq 0.75$, (c) jets p_{T} and (d) leading jet p_{T} for the SM $t\bar{t}$ sample.

7.9 Summary

The large number of pile-up interactions per bunch crossing makes difficult the identification of jets produced in the hard-scatter interaction. A cut on the JVF variable can be used to reject jets from pile-up in the event. The results of performances studies presented in this chapter show that

a cut of 0.75 in the JVF variable gives the optimal performance, i.e. an 80% of rejection against pile-up jets for a 92% hard-scatter jet selection efficiency for a simulated semileptonic $t\bar{t}$ sample. Unfortunately, the JVF variable is not perfectly modeled in the Monte Carlo simulation, as the latter predicts jets with a higher contribution from pile-up than what is seen in data. This causes a difference in the JVF requirement performance in data and simulation and leads to the need of using scale factors in order to match the JVF cut efficiency in observed in simulation to the one in data. These scale factors were derived using a tag & probe method using $Z(\rightarrow \mu\mu/ee)$+jets events with specific selections to obtain a sample of hard-scattering jets and of pile-up jets. After applying them to the simulated samples in the semileptonic $t\bar{t}$ analysis the agreement between data and simulation is well recovered. The associated systematic uncertainties were also presented in this chapter. Variations of the order of 2.5-3% are observed.

Conclusion

In the course of this thesis three main topics have been treated: the performance and validation of the Global Sequential (GS) jet calibration, the search for resonances in lepton plus jets $t\bar{t}$ events and the performance of the Jet Vertex Fraction (JVF) algorithm in lepton plus jets $t\bar{t}$ events with the ATLAS experiment.

The determination of the jet energy scale, its uncertainty and the achievement of an optimal resolution and low flavor sensitivity are major tasks in the ATLAS collaboration to improve the precision of physics analyses with jets in their final states. The GS calibration is an extension of the current jet calibration scheme used in ATLAS called EM+JES. It is derived in Monte Carlo and performs better than the EM+JES scheme in terms of energy resolution and flavor sensitivity. In Chapter 5, the performance of the GS in Monte Carlo, its validation in data using about 35 pb^{-1} of data collected by the ATLAS experiment in 2010 and the evaluation of the associated systematic uncertainty are presented. The systematic uncertainty on the GS calibration is estimated in an inclusive multi-jets sample using a method based on the differences between data and Monte Carlo of the jet properties used as input to the calibration. The final systematic uncertainty is found to be lower than 1% for $|\eta| < 2.8$ and $20 < p_{\text{T}}^{\text{jet}} < 800$ GeV. These results are supported by the results obtained on the γ+jets events in a p_{T} range between 20 GeV and 260 GeV. This uncertainty has to be added in quadrature to the EM+JES calibration systematic uncertainty. Extra systematic uncertainties need to be added to account for the dependence of the jet response on the jet topology and flavour in the future. A brief discussion about the advantages and disadvantages of GS with respect to other jet calibration schemes used in ATLAS is presented in this document.

A search for $t\bar{t}$ resonances has been carried on. The search is done in final states containing at least one electron or muon in a data sample corresponding to an integrated luminosity of 2.05 fb^{-1} collected with the ATLAS experiment. To find a $t\bar{t}$ resonance a serie of cuts are applied to enhance the $t\bar{t}$ topology. The selection used in the analysis is designed for the particular resolved topology. No evidence for a resonance is been found and limits are set on the cross-section times branching ratio for a narrow Z' and a wide Randall Sundrum (RS) Kaluza-Klein (KK) gluon resonances. For the Z' resonance the 95% C.L. observed upper limit range from 9.3 pb for a mass of 500 GeV to 0.95 pb for a mass of 1300 GeV, excluding a leptophobic topcolor Z' boson with 500 GeV$< m_{Z'} <$880 GeV. KK gluons in the RS model with masses between 500 and 1130 GeV are excluded at 95% C.L. These results are part of the paper in Ref. [4].

Further studies can be done to improve the reconstruction efficiency of the $t\bar{t}$ pair as using an hybrid between the dRmin and the χ^2 method or implementing a multivariate analysis. On the other hand, in order to improve the $t\bar{t}$ mass resolution we could think in including the final state radiation jets (if any) in the $t\bar{t}$ pair reconstruction, since the energy carried by the gluons leads to a left tail in the invariant mass distribution. The dominant sources of shape systematic uncertainties in the analysis are the $b-$tagging efficiency with $^{+18.9\%}_{-16.5\%}$ ($^{+12.7\%}_{-11.2\%}$) event yield variation for the Z' of 1 TeV (background), the jet energy scale with $^{+3.4\%}_{-4.2\%}$ ($^{+14.5\%}_{-11.6\%}$) and the modeling of initial state radiation with $^{+0.9\%}_{-8.5\%}$ ($^{+4.3\%}_{-10.4\%}$). Reduction of the systematic uncertainties is very important. In order to do that a better understanding of the ISR/FSR and jet performance are needed, as well as more precise studies for the determination of background estimates. The combination of this analysis with the results obtained from the search for $t\bar{t}$ resonances in events with highly boosted tops will allow to gain more sensitivity at high $t\bar{t}$ pair mass.

Related to this search and in order to improve the efficiency of the $t\bar{t}$ pair reconstruction

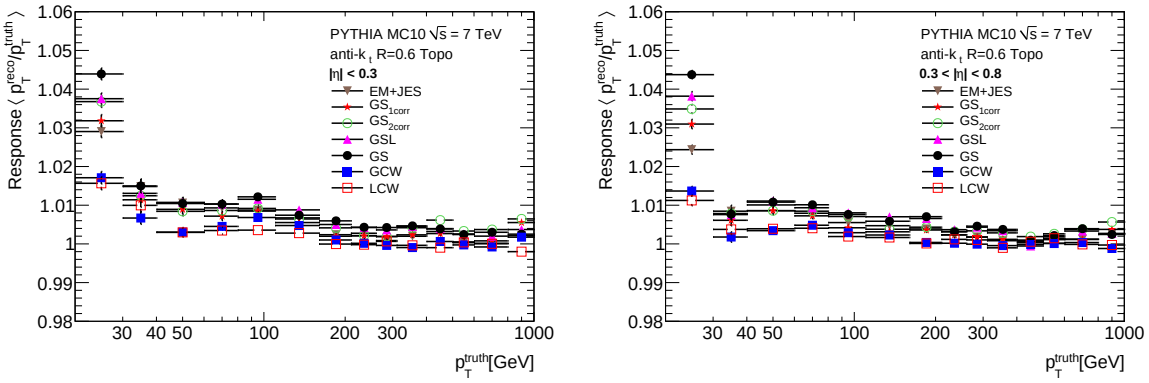
and selection at high luminosities, performance studies of the jet vertex fraction (JVF) in top-quark pairs topologies have been also performed. JVF is a variable calculated for each jet that quantifies the fraction of track transverse momentum associated to the jet from the hard-scattering interaction. Therefore, a cut in the JVF variable can help to filter jets coming from pile-up interactions. The optimal JVF cut for semileptonic top-quark analyses is found to be 0.75. Scale factors to account for differences in the performance of the JVF cut in data and Monte Carlo are calculated, as well as their associated systematic uncertainties. Variations of the order of 2.5-3% are observed for the systematic uncertainties. The results derived in this chapter are being used in most of the top analyses with the whole dataset collected in 2011, corresponding to an integrated luminosity of 4.7 fb^{-1}. Further studies could be done to complement the results presented in this chapter. The derivation of bidimensional scale factors is one of them, to correct for the small dependence observed as a function of η. Although, the pile-up jet contamination fraction observed in the hard-scatter enriched region is small, it could be removed in order to get more accurate scale factors for the hard-scatter jets. In addition, the derivation of the scale factors for pile-up jets with more statistic and further studies to define a better pile-up enriched region are very important tasks that can be done in future studies.

El viaje no termina jamás... El objetivo de un viaje es solo el inicio de otro viaje..
José Saramago, 1981

Appendix A

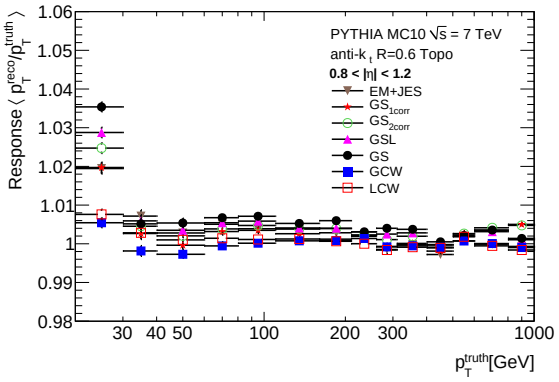
GS calibration: jet response linearity and jet resolution

This section contains the jet response and jet resolution plots for the GS scheme summary plots of Chapter 5 Section 5.4.1. For jets reconstructed with the anti-k_t algorithm with $R = 0.6$ using topoclusters as inputs and for all η regions.



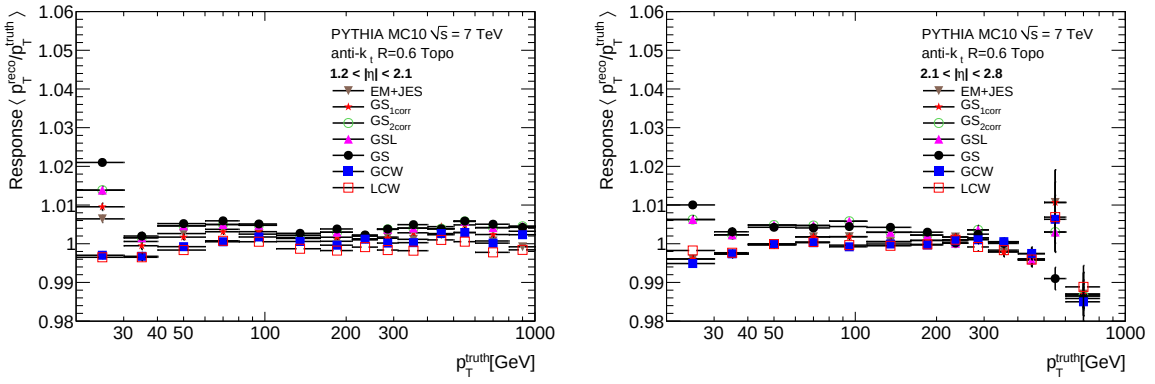
(a)

(b)



(c)

Figure A.1: Average jet response as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ for $|\eta| < 1.2$ after each GS correction in the nominal Pythia simulation for jets reconstructed with the anti-k_t algorithm $R = 0.6$. GCW and LCW results are also shown for comparison.



(a) (b)

(c) (d)

Figure A.2: Average jet response as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ for $1.2 < |\eta| < 4.5$ after each GS correction in the nominal PYTHIA simulation for jets reconstructed with the anti-k_t algorithm $R = 0.6$. GCW and LCW results are also shown for comparison.

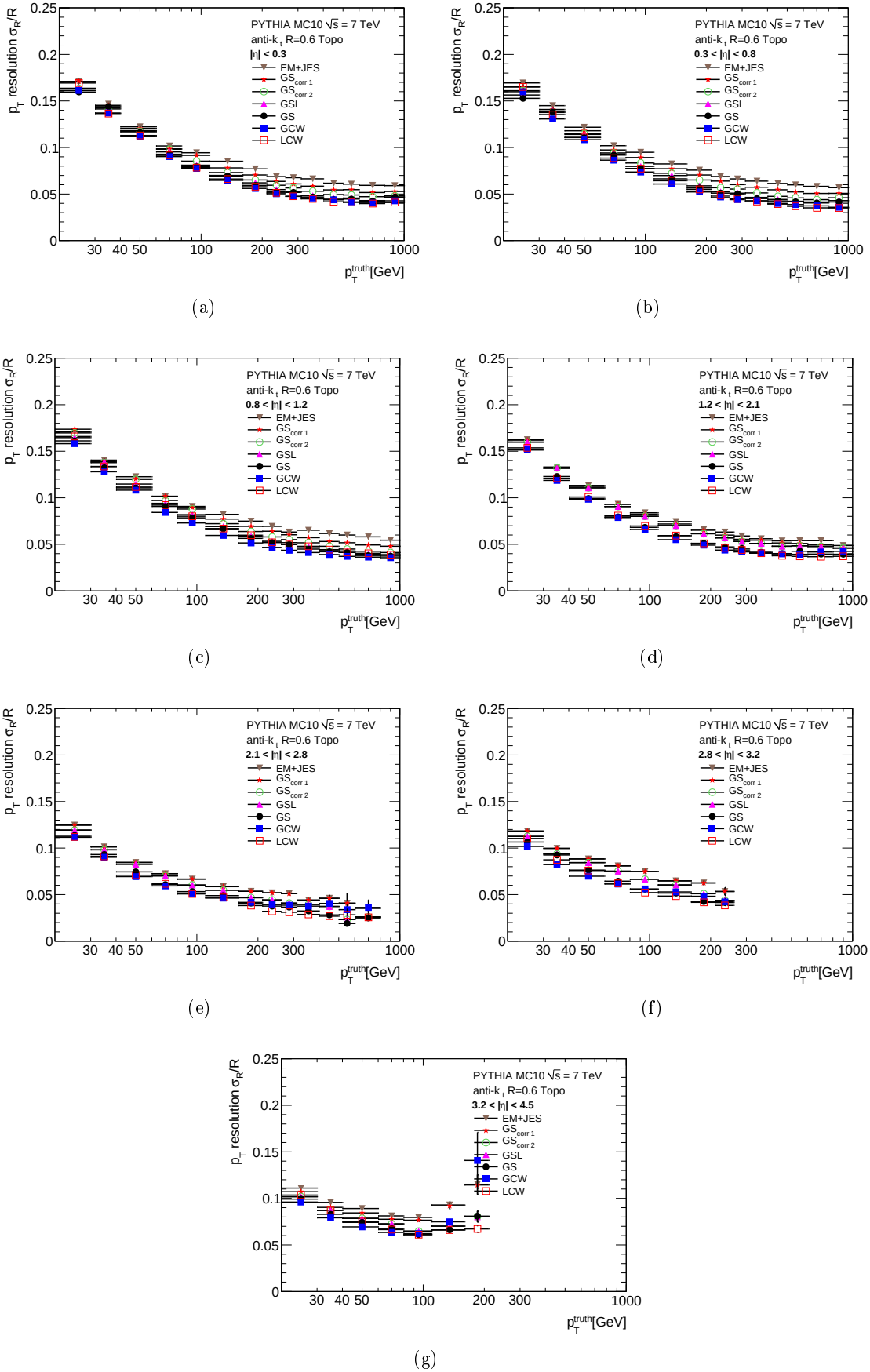
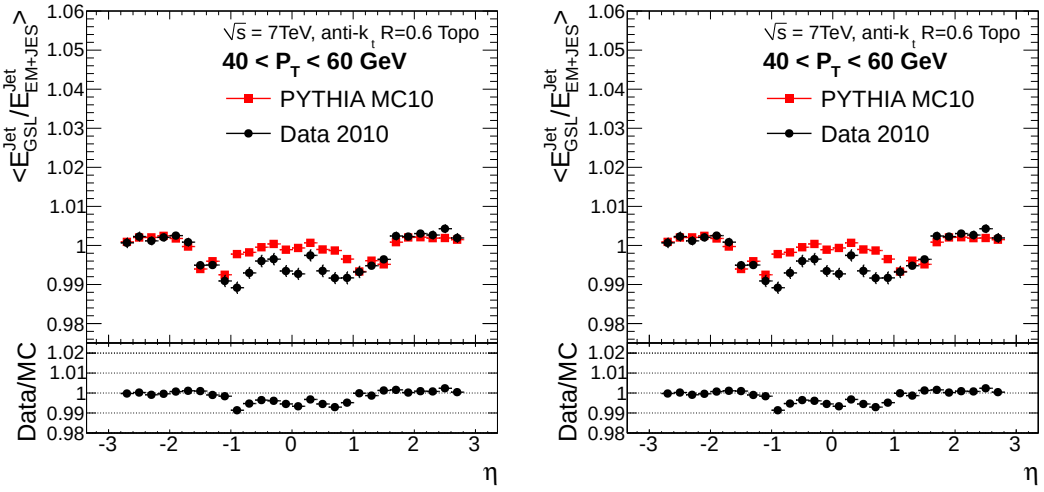


Figure A.3: Average fractional jet resolution as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ for different $|\eta|$ regions after each GS correction in the nominal PYTHIA simulation for jets reconstructed with the anti-k_t algorithm $R = 0.6$. GCW and LCW results are also shown for comparison.

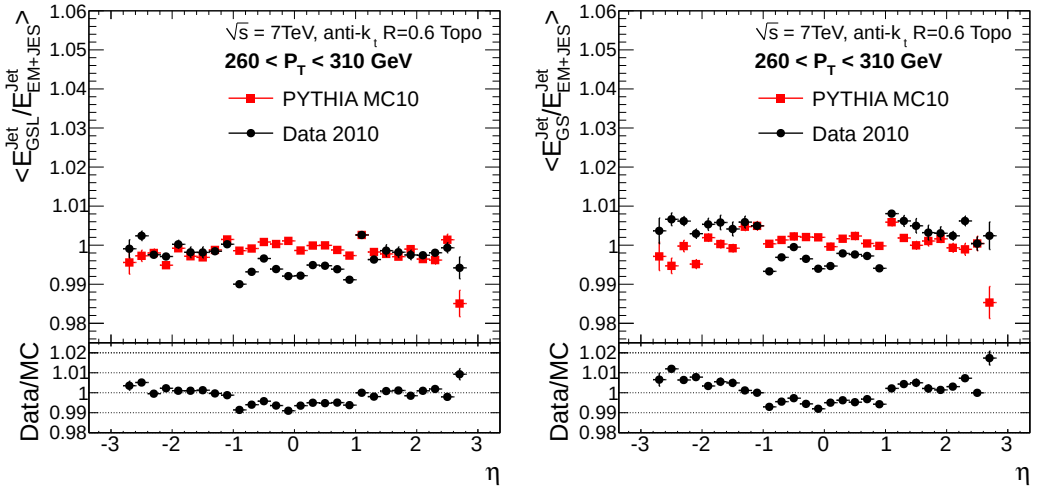
Appendix B

GS calibration: Systematic uncertainty plots

This section contains the GS(L) systematic uncertainty summary plots of Chapter 5 Section 5.6.1 for all η regions.

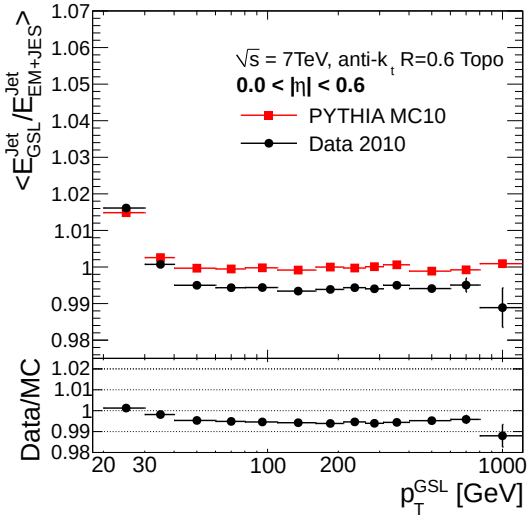


(a) (b)

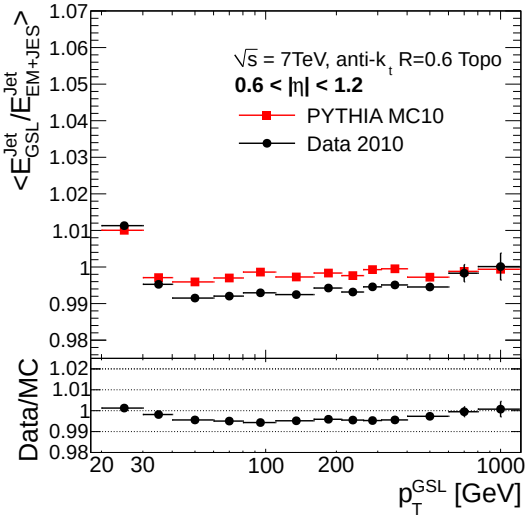


(c) (d)

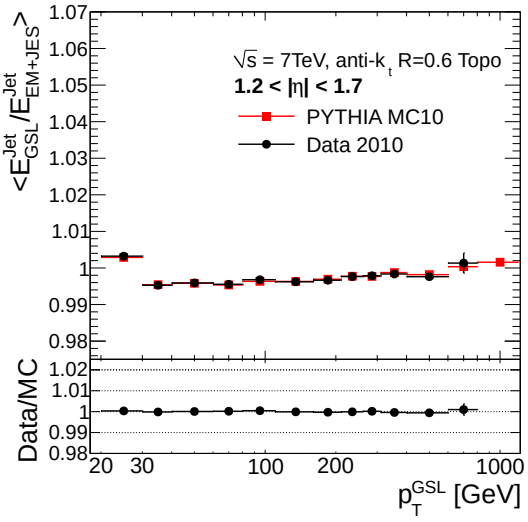
Figure B.1: Average jet energy after GSL (left) and GS (right) divided by the average jet energy after the EM+JES calibration as a function of η in data and the reference Pythia sample for $40 < p_{\mathrm{T}}^{\mathrm{jet}} < 80$ GeV (upper part) and $260 < p_{\mathrm{T}}^{\mathrm{jet}} < 3100$ GeV (lower part). The double ratio $(E_{\mathrm{GS(L)}}/E_{\mathrm{EM+JES}})_{\mathrm{data}}/(E_{\mathrm{GS(L)}}/E_{\mathrm{EM+JES}})_{\mathrm{MC}}$ is shown at the bottom of each plot.



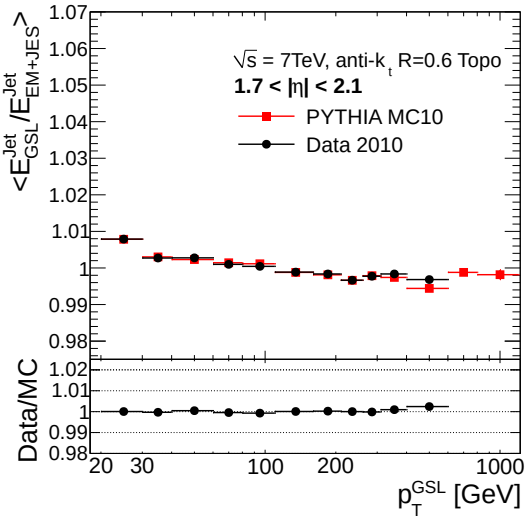
(a)



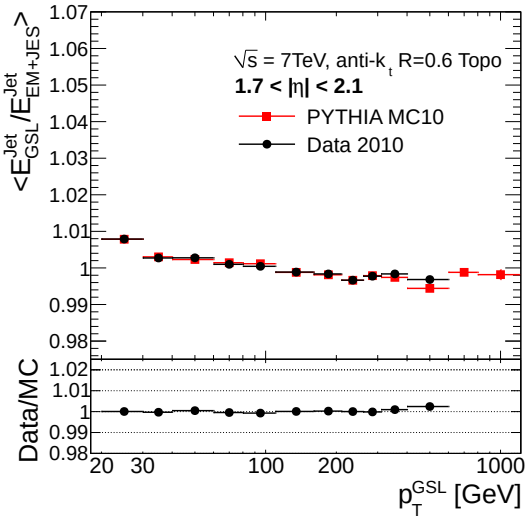
(b)



(c)

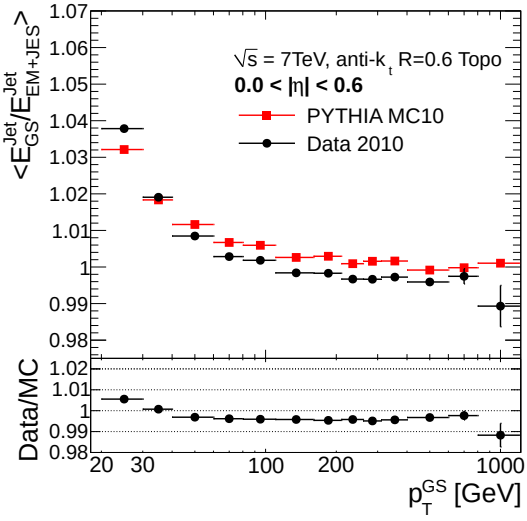


(d)

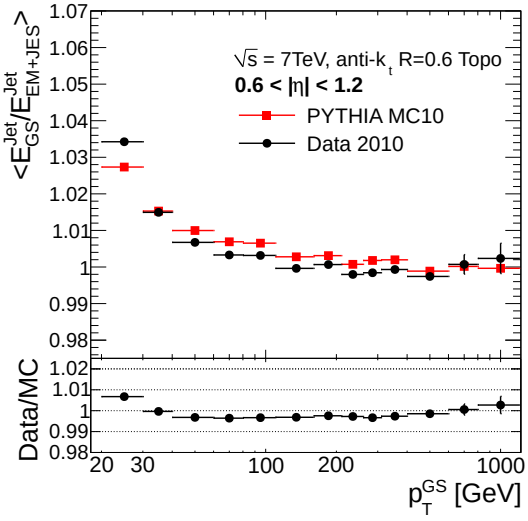


(e)

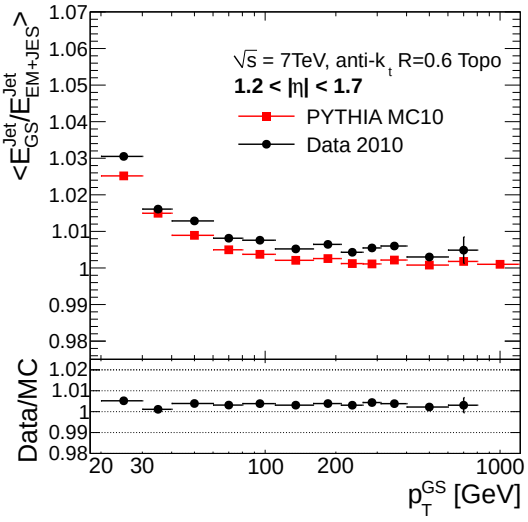
Figure B.2: Average jet energy after GSL divided by the average jet energy after the EM+JES calibration as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ in data and the reference PYTHIA sample in the barrel. The double ratio $(E_{\mathrm{GS(L)}}/E_{\mathrm{EM+JES}})_{\mathrm{data}}/(E_{\mathrm{GS(L)}}/E_{\mathrm{EM+JES}})_{\mathrm{MC}}$ is shown at the bottom.



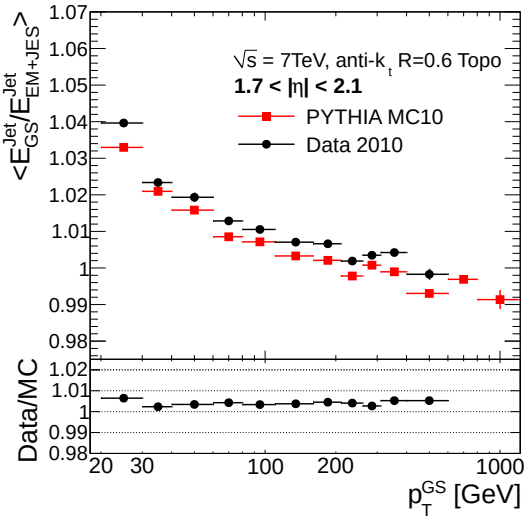
(a)



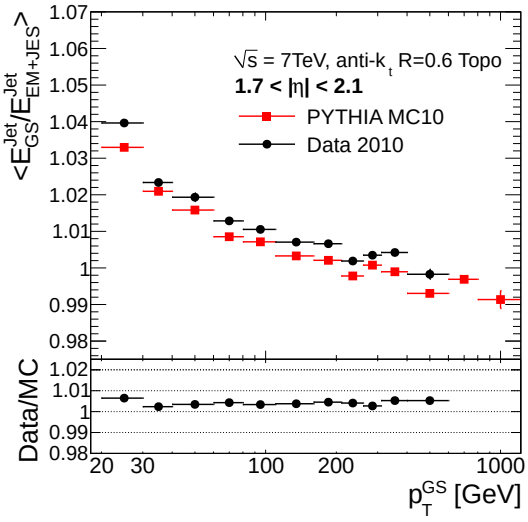
(b)



(c)



(d)



(e)

Figure B.3: Average jet energy after GS divided by the average jet energy after the EM+JES calibration as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ in data and the reference PYTHIA sample in the barrel. The double ratio $(E_{\mathrm{GS(L)}}/E_{\mathrm{EM+JES}})_{\mathrm{data}}/(E_{\mathrm{GS(L)}}/E_{\mathrm{EM+JES}})_{\mathrm{MC}}$ is shown at the bottom.

Appendix C

GS calibration: Statistical uncertainties on data-base GS corrections

The mean value of the response at the GS scale in a given p_T and η bins correspond to:

$$< R^{GS}(p_T, \eta, x_i) >= \frac{\sum_{jets:j} R^{GS}(p_T, \eta, x_i^j)}{N_{jets}} \tag{C.1}$$

where N_{jets} is the number of jets in that given p_T and η bins. And x_i is the corresponding set of variable used to correct the response: layer energy fractions and/or jet width. We also make N_{bins} bins in the variable x_i, then we can express equation (1) as:

$$< R^{GS}(p_T, \eta, x_i) >= \frac{\sum_{xbin:k} N_k R^{GS}(p_T, \eta, x_i^k)}{N_{jets}} \tag{C.2}$$

where N_k correspond to the number of jets in bin k.

Using equation (2), we can express the statistical error in a given p_T and η bin as:

$$\sigma^2_{<R^{GS}(p_T,\eta,x_i)>} = \sum_k \Big(\frac{\partial < R^{GS}(p_T, \eta, x_i) >}{\partial R^{GS}(p_T, \eta, x_i^k)} \Big)^2 \sigma^2_{R^{GS}(p_T,\eta,x_i^k)} = \sum_k \Big(\frac{N_k}{N_{jets}} \Big)^2 \sigma^2_{R^{GS}(p_T,\eta,x_i^k)} \tag{C.3}$$

In order to decide the binning in the variable x_i we use a dynamical binning procedure which allows, in principle, to have the same number of entries in each k bin (i.e. $N_k = constant = c$). This means that in a good aproximation all the $\sigma^2_{R^{GS}(p_T,\eta,x_i^k)}$ are the same. Then:

$$\sigma^2_{<R^{GS}(p_T,\eta,x_i)>} = \sigma^2_{R^{GS}(p_T,\eta,x_i^k)} \frac{\sum_k N_k^2}{N_{jets}^2} \tag{C.4}$$

where:

$$\frac{\sum_k N_k^2}{N_{jets}^2} = \frac{\sum_k N_k^2}{(\sum_k N_k)^2} = \frac{\sum_k c^2}{(\sum_k c)^2} = \frac{N_{bins} c^2}{(N_{bins} c)^2} = \frac{1}{N_{bins}} \tag{C.5}$$

Using equation (5) in (4) we get:

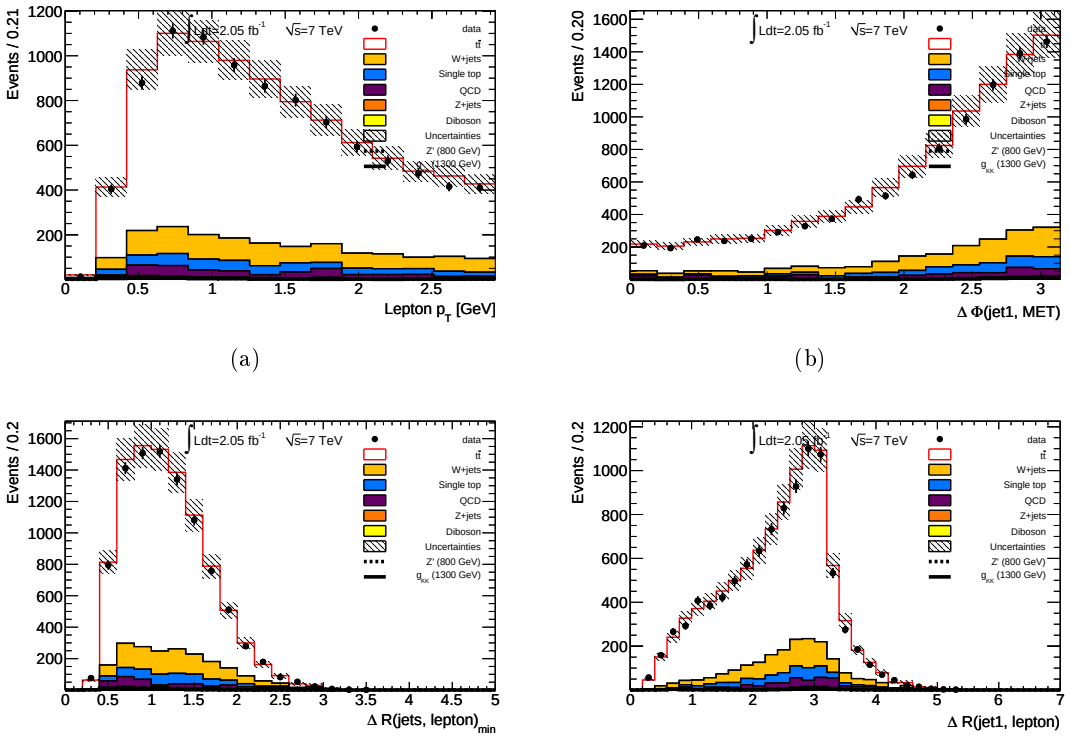
$$\sigma_{<R^{GS}(p_T,\eta,x_i)>} = \sigma_{R^{GS}(p_T,\eta,x_i^k)} \frac{1}{\sqrt{N_{bins}}} \tag{C.6}$$

Appendix D

Search for $t\bar{t}$ resonances in ATLAS: Electron channel

Figures D.2, D.1 and D.3 show the distributiona for some kinematic variables for data, the different backgrounds considered in the analysis and two signal points, a Z' with $m_{Z'} = 800$ GeV and a KK gluon with $m_{g_{KK}} = 1300$ GeV in the electron channel. Figure D.4 shows the agreement between data and expectation from the sum of all backgrounds in the $t\bar{t}$ mass spectra. Figure D.5 shows the relative difference in reconstructed $t\bar{t}$ mass between data and expectation for the different methods in the electron channel.

Figures D.6 and D.7 show the impact of the individual shape systematics on the dominant $t\bar{t}$ and W+jets backgrounds on the reconstructed $t\bar{t}$ mass spectra for the electron channel. The algorithm used to reconstruct the $t\bar{t}$ mass is the dRmin method (see Section 6.9.3).



(a) (b)

(c) (d)

Figure D.1: (a) $\Delta\phi$ between the electron and the $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$, (b) $\Delta\phi$ between the leading jet and the $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(jet1, E_{\mathrm{T}}^{\mathrm{miss}})$, (c) minimum ΔR between jets in the event and the electron, $\Delta R_{min}(\ell, jets)$, and (d) ΔR between the leading jet and the electron, $\Delta R(\ell, jet1)$ in the electron channel. The grey hashed area shows the total background normalization uncertainty.

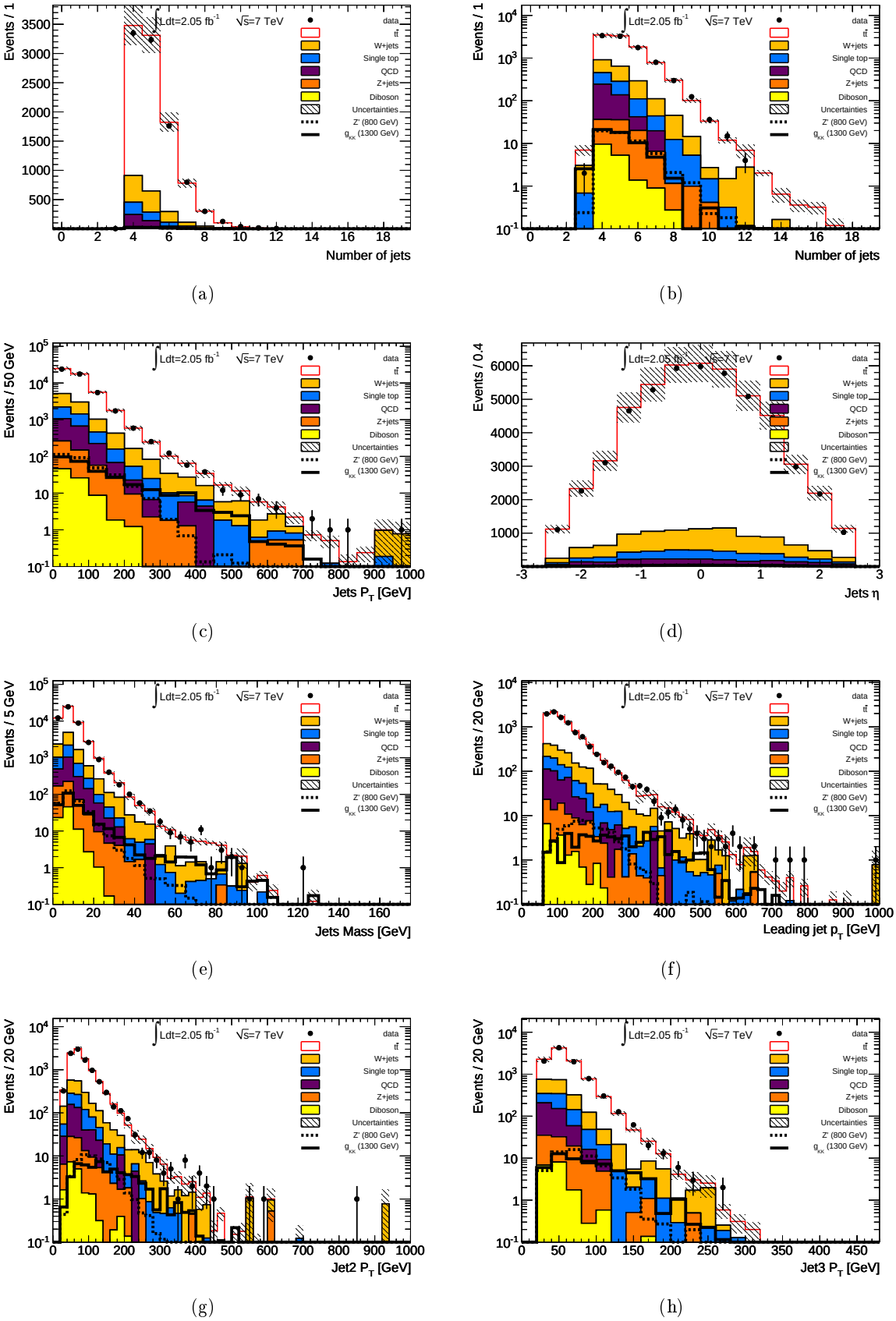
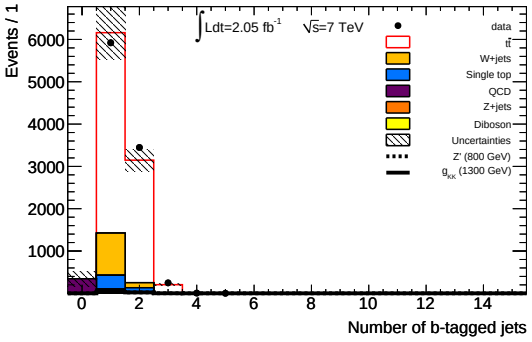
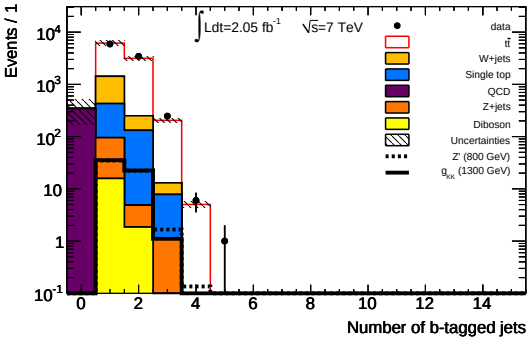


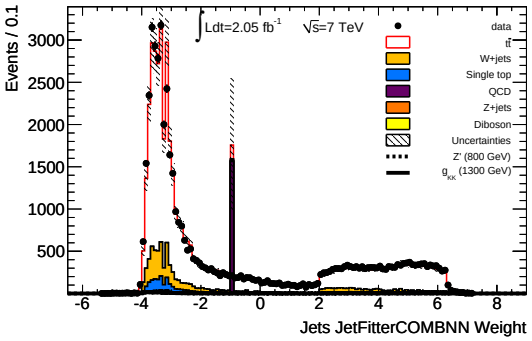
Figure D.2: Number of jets with $p_{\mathrm{T}} > 20$ GeV in (a) linear and (b) logarithmic scales, jets (c) p_{T}, (d) η and (e) mass distributions, (f) leading jet p_{T}, (g) second jet p_{T} and (h) third jet p_{T} after all selection criteria were applied in the electron channel. The grey hashed area shows the total background normalization uncertainty.



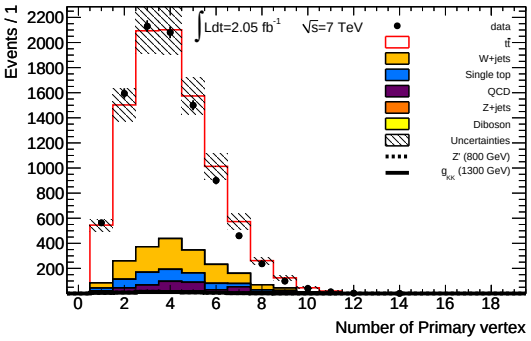
(a)



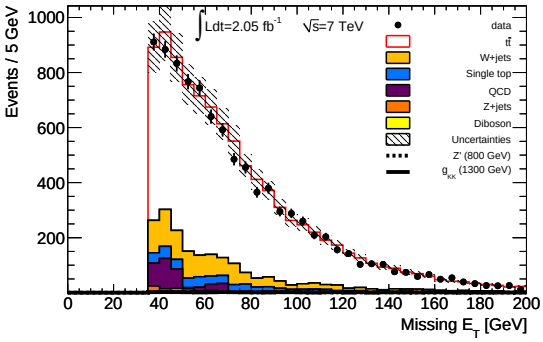
(b)



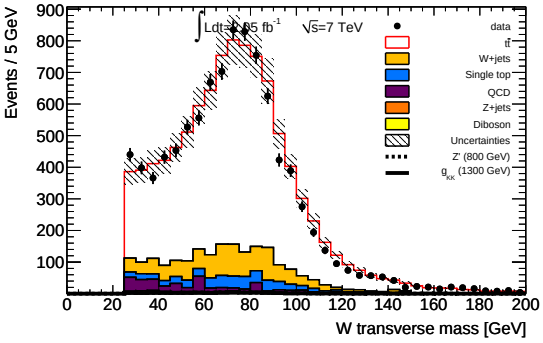
(c)



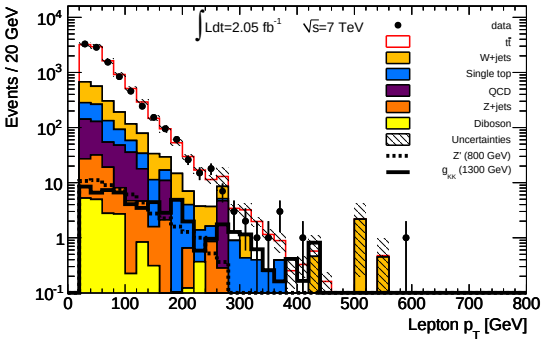
(d)



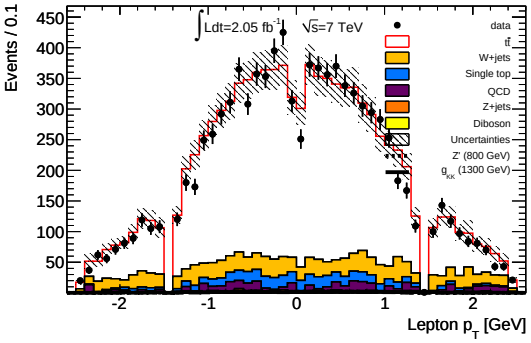
(e)



(f)



(g)



(h)

Figure D.3: Number of $b-$tagged jets with $p_{\mathrm{T}} > 20$ GeV in (a) linear and (b) logarithmic scales, (c) JetFitterCOMBNN weight for all jets, (d) number of primary vertices, electron (e) p_{T} and (f) η, (g) $E_{\mathrm{T}}^{\mathrm{miss}}$ and (h) W boson transverse mass after all selection criteria were applied in the electron channel. The grey hashed area shows the total background normalization uncertainty.

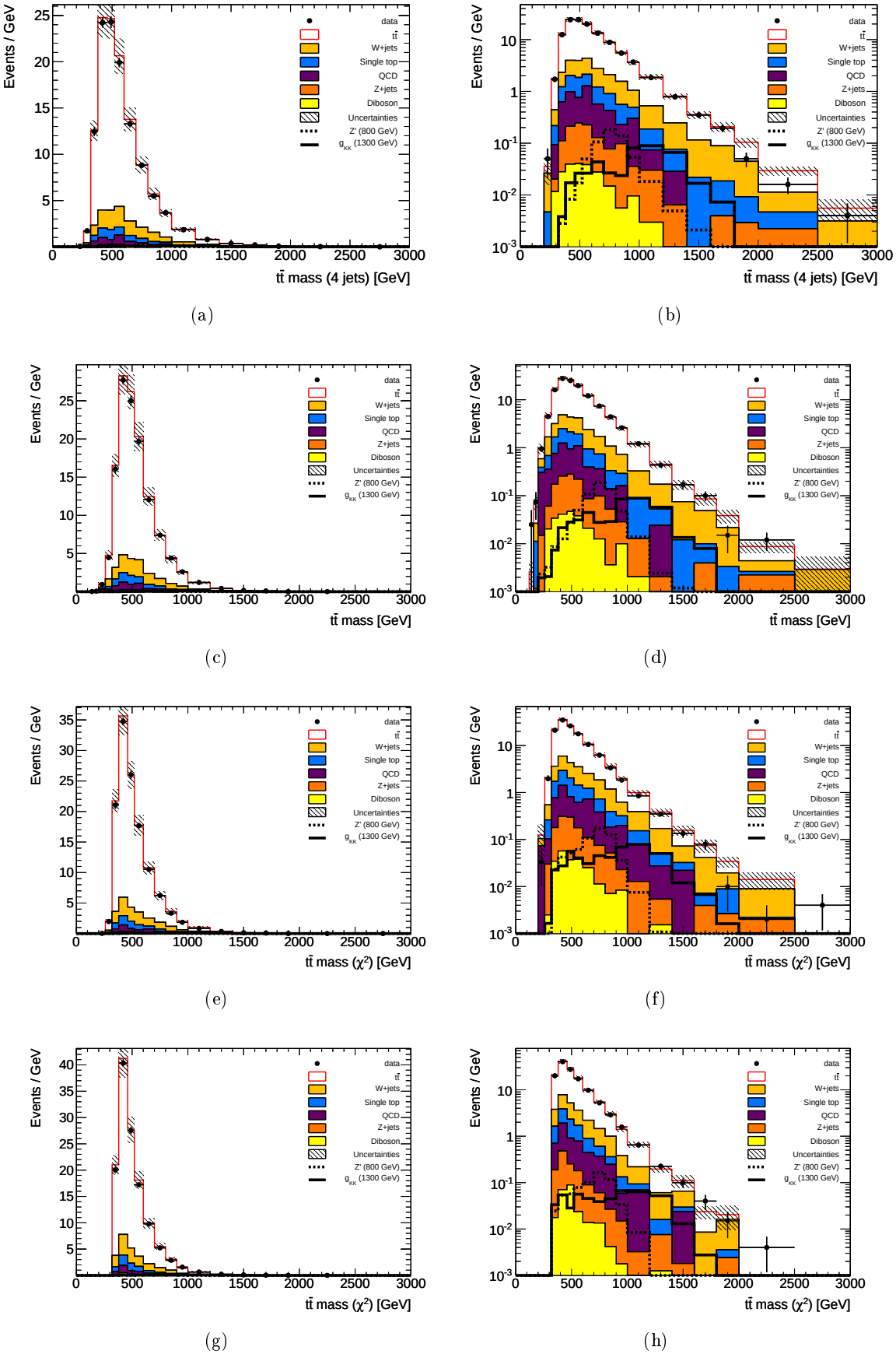
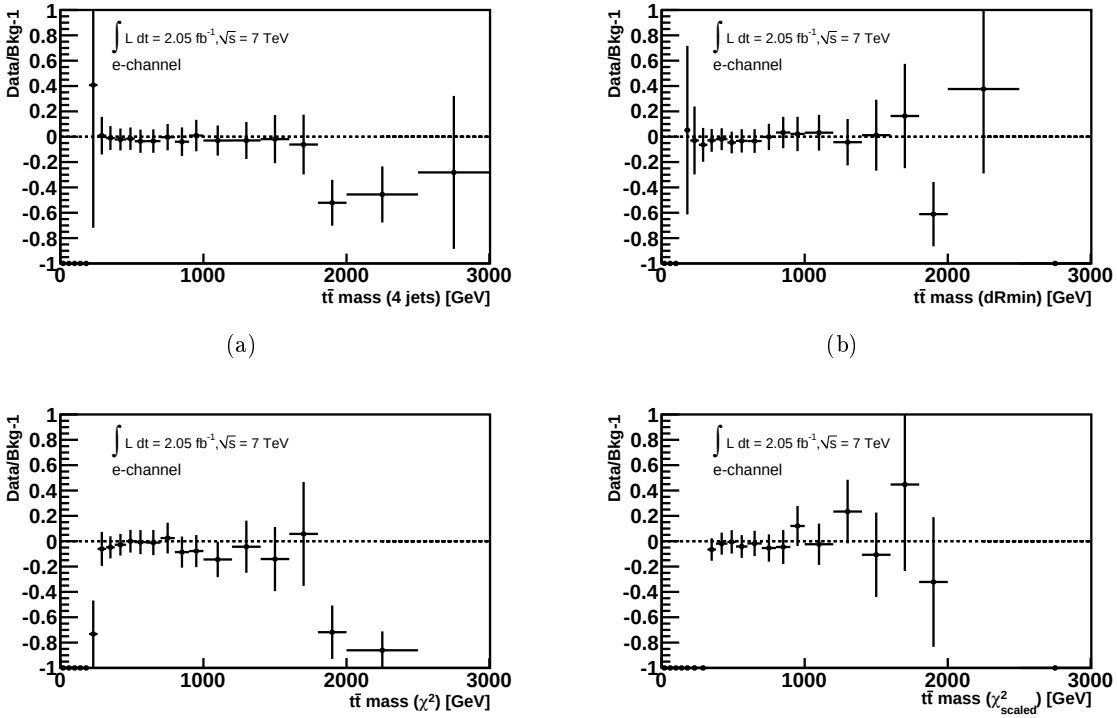


Figure D.4: Reconstructed $t\bar{t}$ mass using the (a,b) four hardest jets, (c,d) dRmin, (e,f) χ^2 and (g,h) χ^2 scaled methods in linear (left) and logarithmic scales (right). The grey hashed area shows the total background normalization uncertainty. Electron channel only.



(a) (b)

(c) (d)

Figure D.5: Relative difference in reconstructed $t\bar{t}$ mass between data and expectation for the (a) four hardest jets, (b) dRmin, (c) χ^2 and (d) χ^2 scaled methods. Electron channel only.

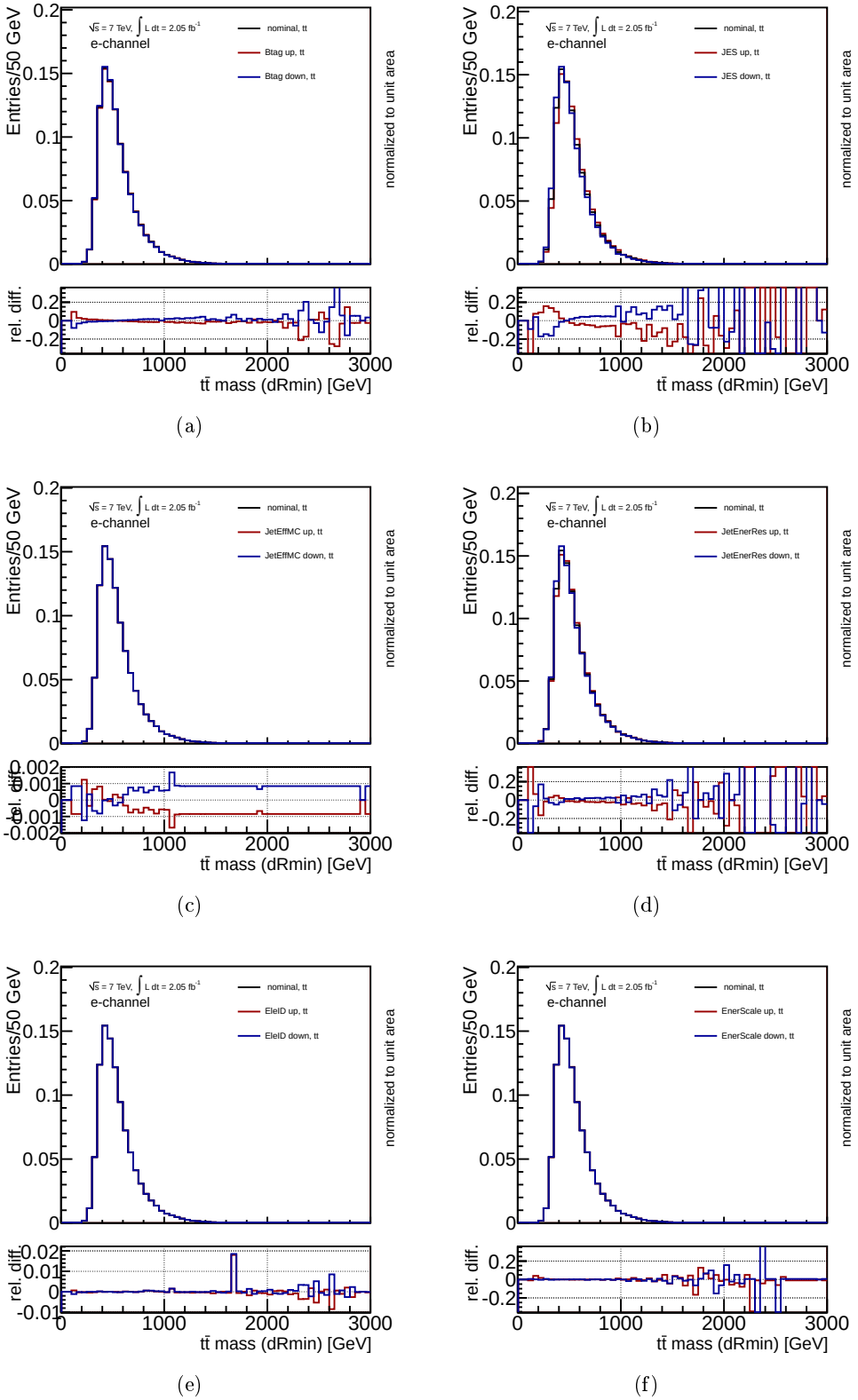


Figure D.6: Impact the (a) b-tagging efficiency, (b) jet energy scale, (c) jet reconstruction efficiency, (d) jet energy resolution, (e) electron identification efficiency and (f) electron energy scale systematic uncertainties on the $t\bar{t}$ mass spectra for the SM $t\bar{t}$ background in the electron channel [8].

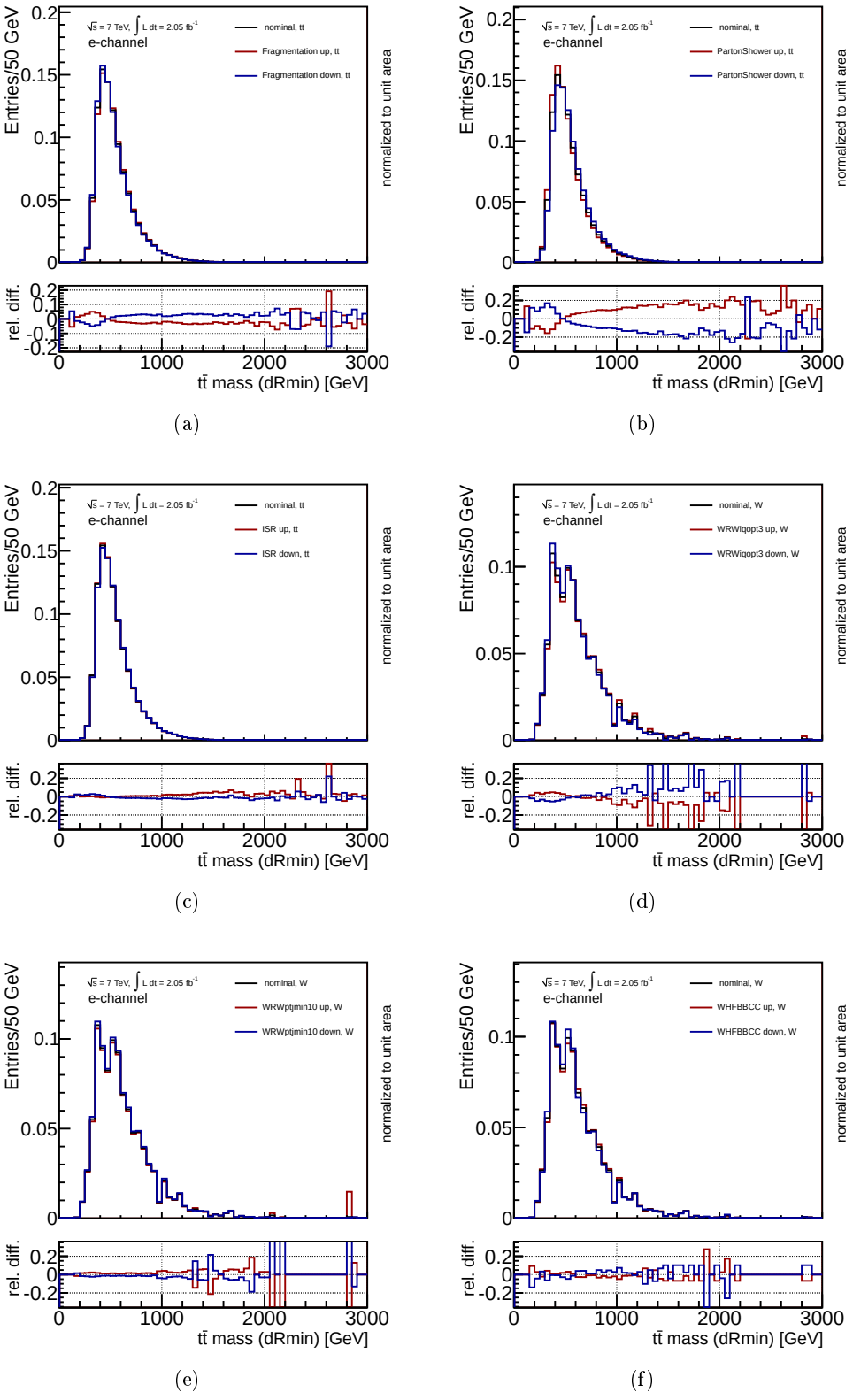


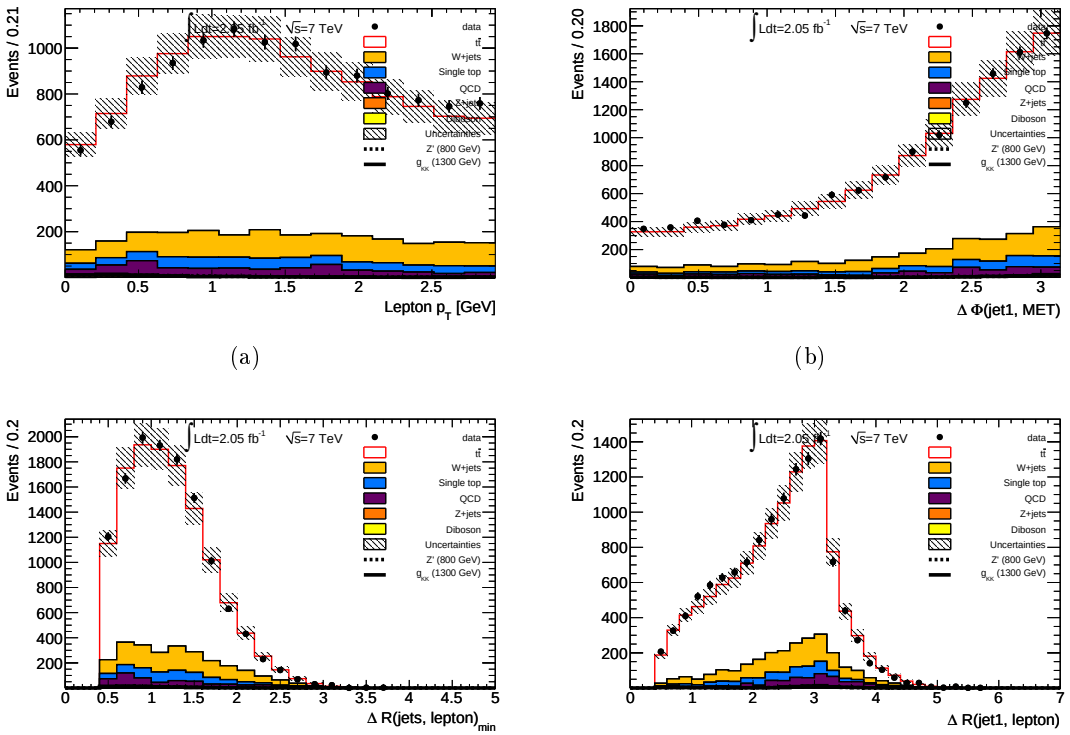
Figure D.7: Impact of the (a) $m_{t\bar{t}}$ shape, (b) parton shower and fragmentation and (c) ISR systematic uncertainties on the $t\bar{t}$ mass spectra for the SM $t\bar{t}$ background in the electron channel. The impact of the (d) W+jets shape (`iqopt3`), (e) W+jets shape (`ptjmin`) and (f) W+jets flavor composition systematic uncertainties on the $t\bar{t}$ mass spectra for the W+jets background in the electron channel [8].

Appendix E

Search for $t\bar{t}$ resonances in ATLAS: Muon channel

Figures E.2, E.1 and E.3 show the distributions for some kinematic variables for data, the different backgrounds considered in the analysis and two signal points, a Z' with $m_{Z'} = 800$ GeV and a KK gluon with $m_{gKK} = 1300$ GeV in the muon channel. Figure E.4 shows the agreement between data and expectation from the sum of all backgrounds in the $t\bar{t}$ mass spectra. Figure E.5 shows the relative difference in reconstructed $t\bar{t}$ mass between data and expectation for the different methods in the muon channel.

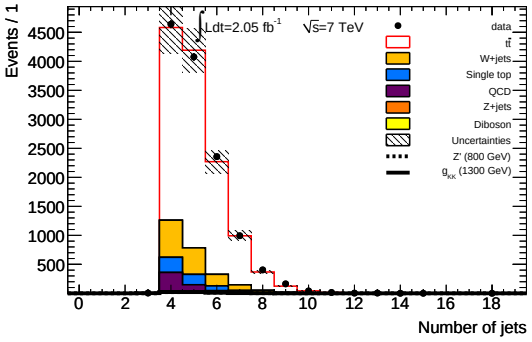
Figures E.6 and E.7 show the impact of the individual shape systematics on the dominant $t\bar{t}$ and W+jets backgrounds on the reconstructed $t\bar{t}$ mass spectra for the muon channel. The algorithm used to reconstruct the $t\bar{t}$ mass is the dRmin method (see Section 6.9.3).



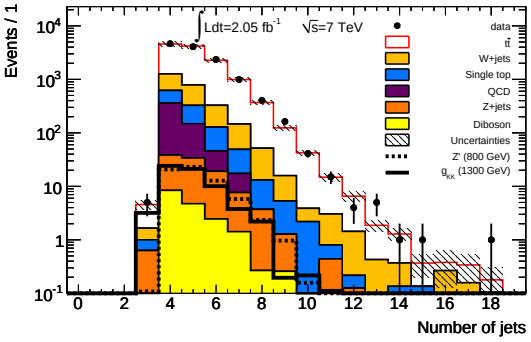
(a) (b)

(c) (d)

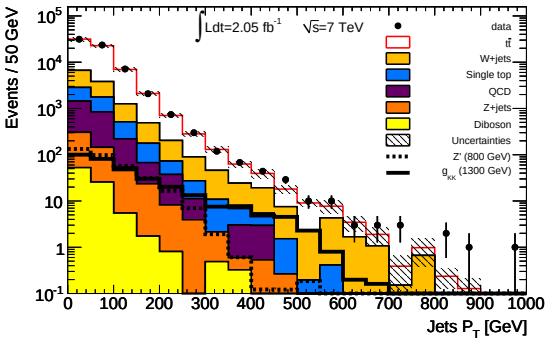
Figure E.1: (a) $\Delta\phi$ between the muon and the $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$, (b) $\Delta\phi$ between the leading jet and the $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi(jet1, E_{\mathrm{T}}^{\mathrm{miss}})$, (c) minimum ΔR between jets in the event and the muon, $\Delta R_{min}(\ell, jets)$, and (d) ΔR between the leading jet and the muon, $\Delta R(\ell, jet1)$ in the muon channel. The grey hashed area shows the total background normalization uncertainty.



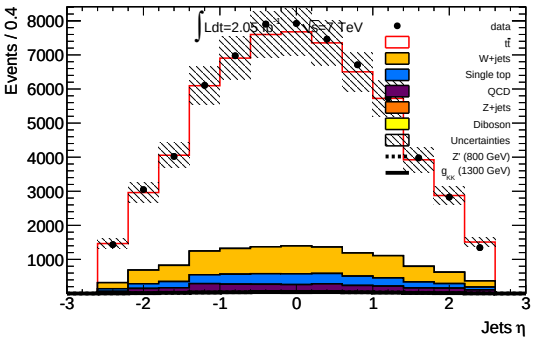
(a)



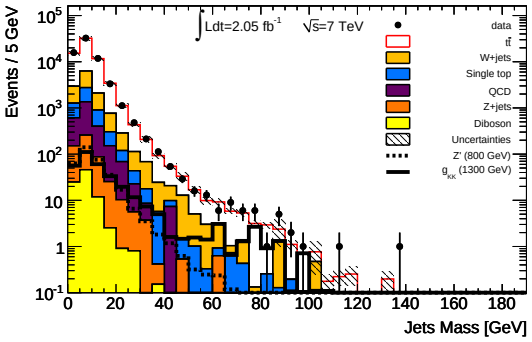
(b)



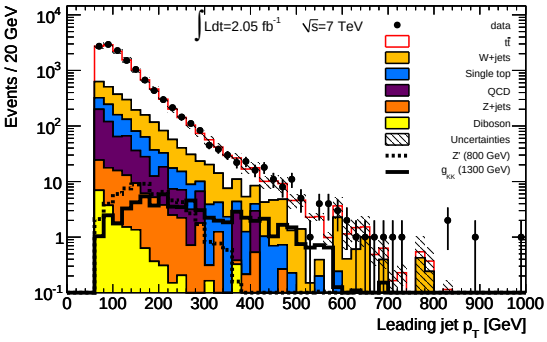
(c)



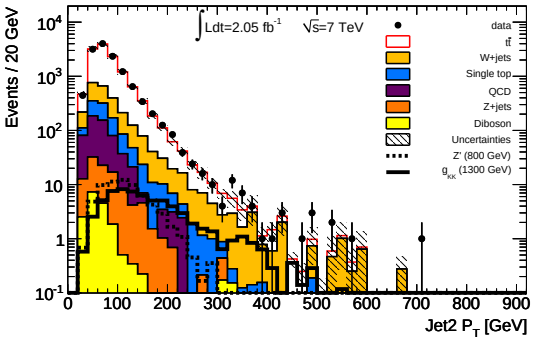
(d)



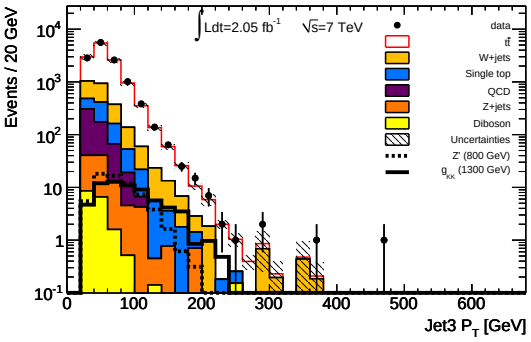
(e)



(f)

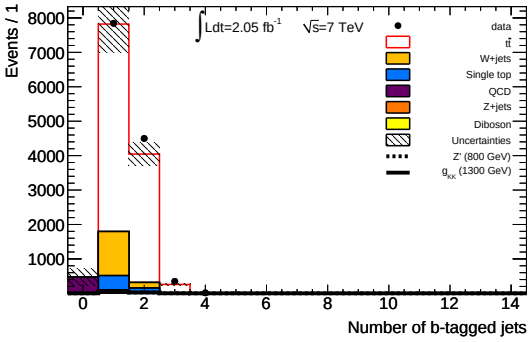


(g)

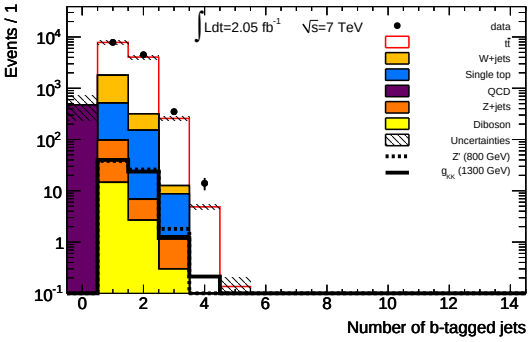


(h)

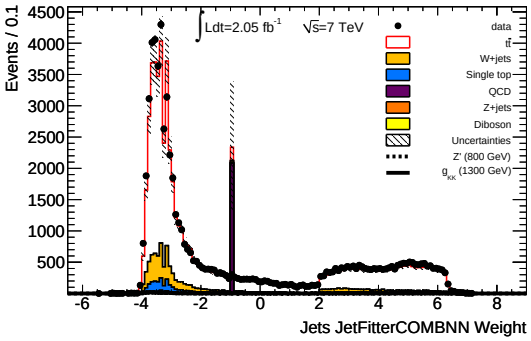
Figure E.2: Number of jets with $p_{\mathrm{T}} > 20$ GeV in (a) linear and (b) logarithmic scales, jets (c) p_{T}, (d) η and (e) mass distributions, (f) leading jet p_{T}, (g) second jet p_{T} and (h) third jet p_{T} after all selection criteria were applied. The grey hashed area shows the total background normalization uncertainty.



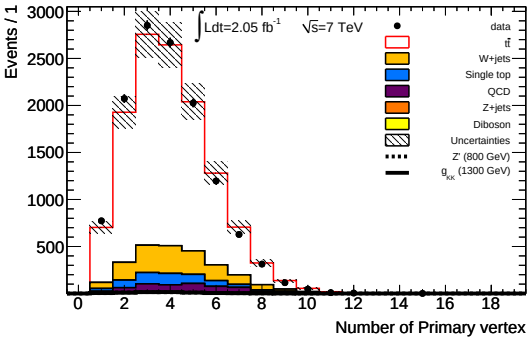
(a)



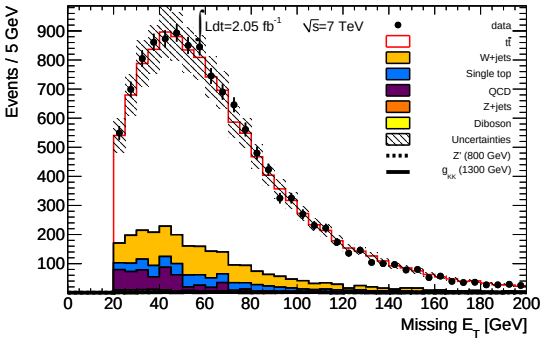
(b)



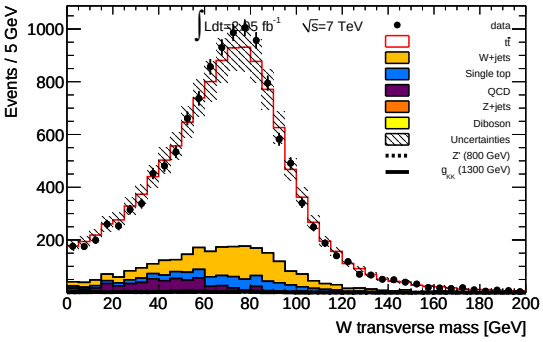
(c)



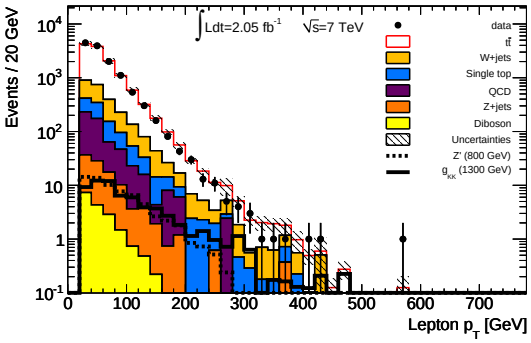
(d)



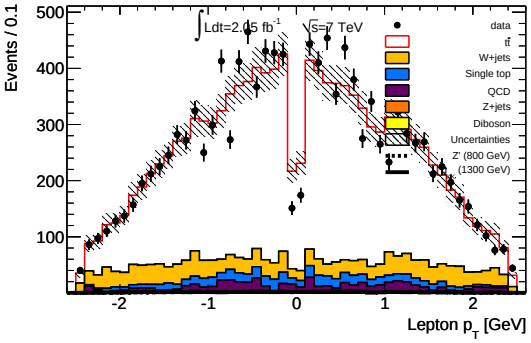
(e)



(f)



(g)



(h)

Figure E.3: Number of $b-$tagged jets with $p_{\mathrm{T}} > 20$ GeV in (a) linear and (b) logarithmic scales, (c) JetFitterCOMBNN weight for all jets, (d) number of primary vertices, muon (e) p_{T} and (f) η, (g) $E_{\mathrm{T}}^{\mathrm{miss}}$ and (h) W boson transverse mass after all selection criteria were applied. The grey hashed area shows the total background normalization uncertainty. Muon channel only.

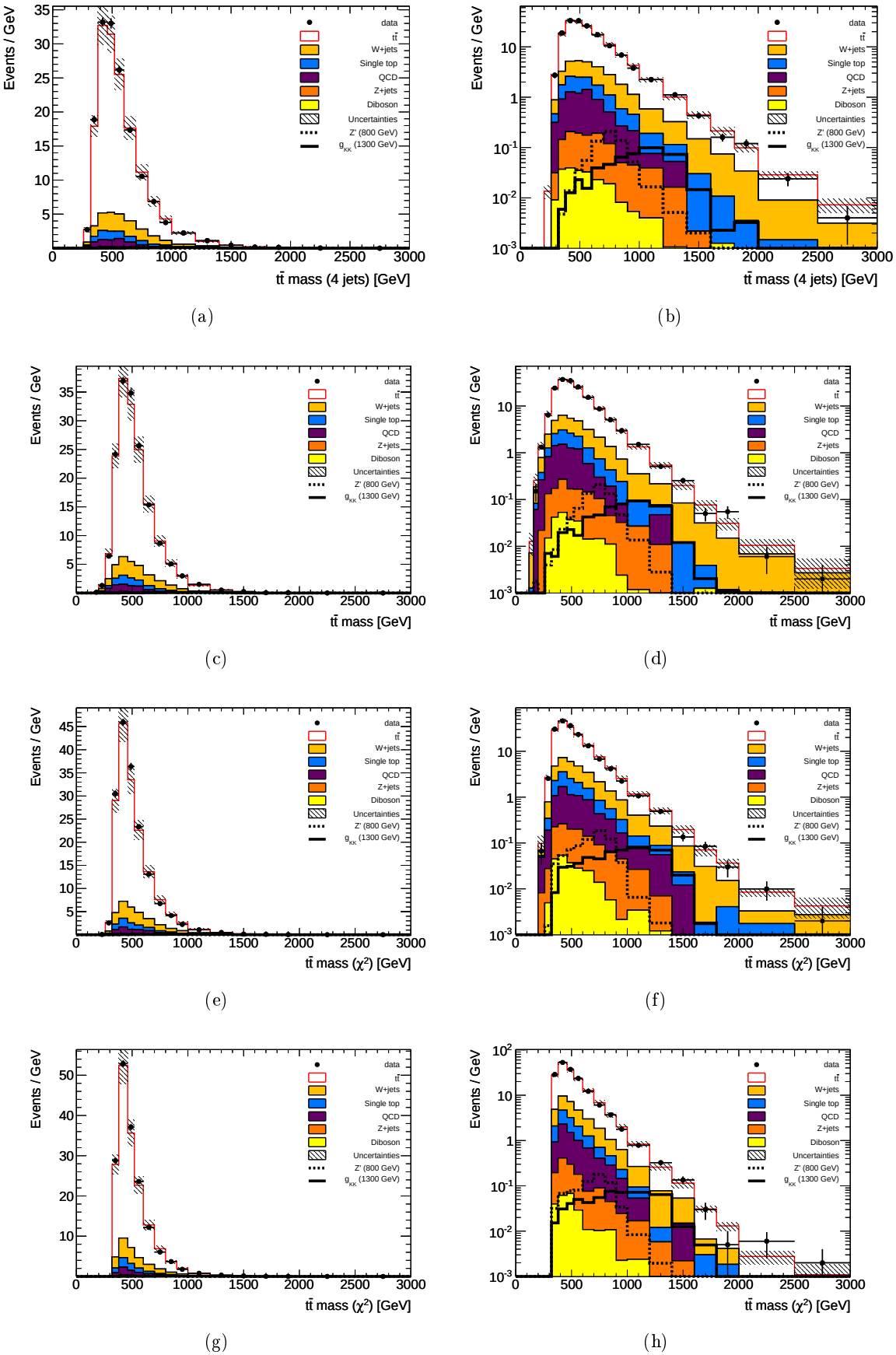
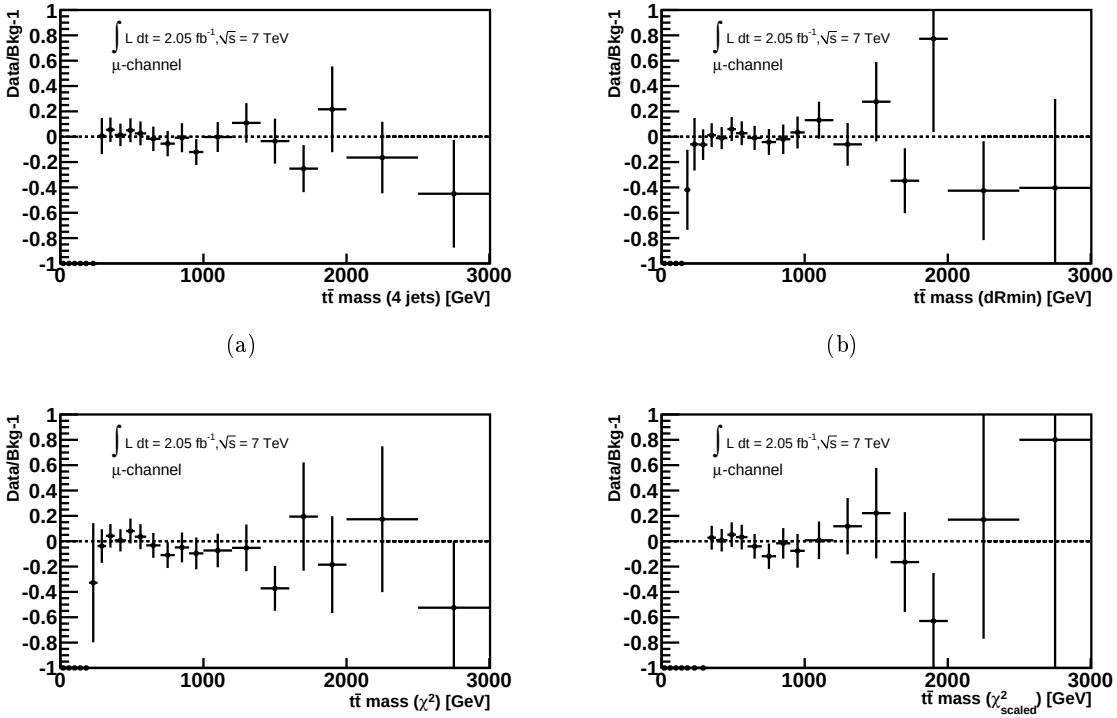


Figure E.4: Reconstructed $t\bar{t}$ mass using the (a,b) four hardest jets, (c,d) dRmin, (e,f) χ^2 and (g,h) χ^2 scaled methods in linear (left) and logarithmic scales (right). The grey hashed area shows the total background normalization uncertainty. Muon channel only.



(a) (b)

(c) (d)

Figure E.5: Relative difference in reconstructed $t\bar{t}$ mass between data and expectation for the (a) four hardest jets, (b) dRmin, (c) χ^2 and (d) χ^2 scaled methods. Muon channel only.

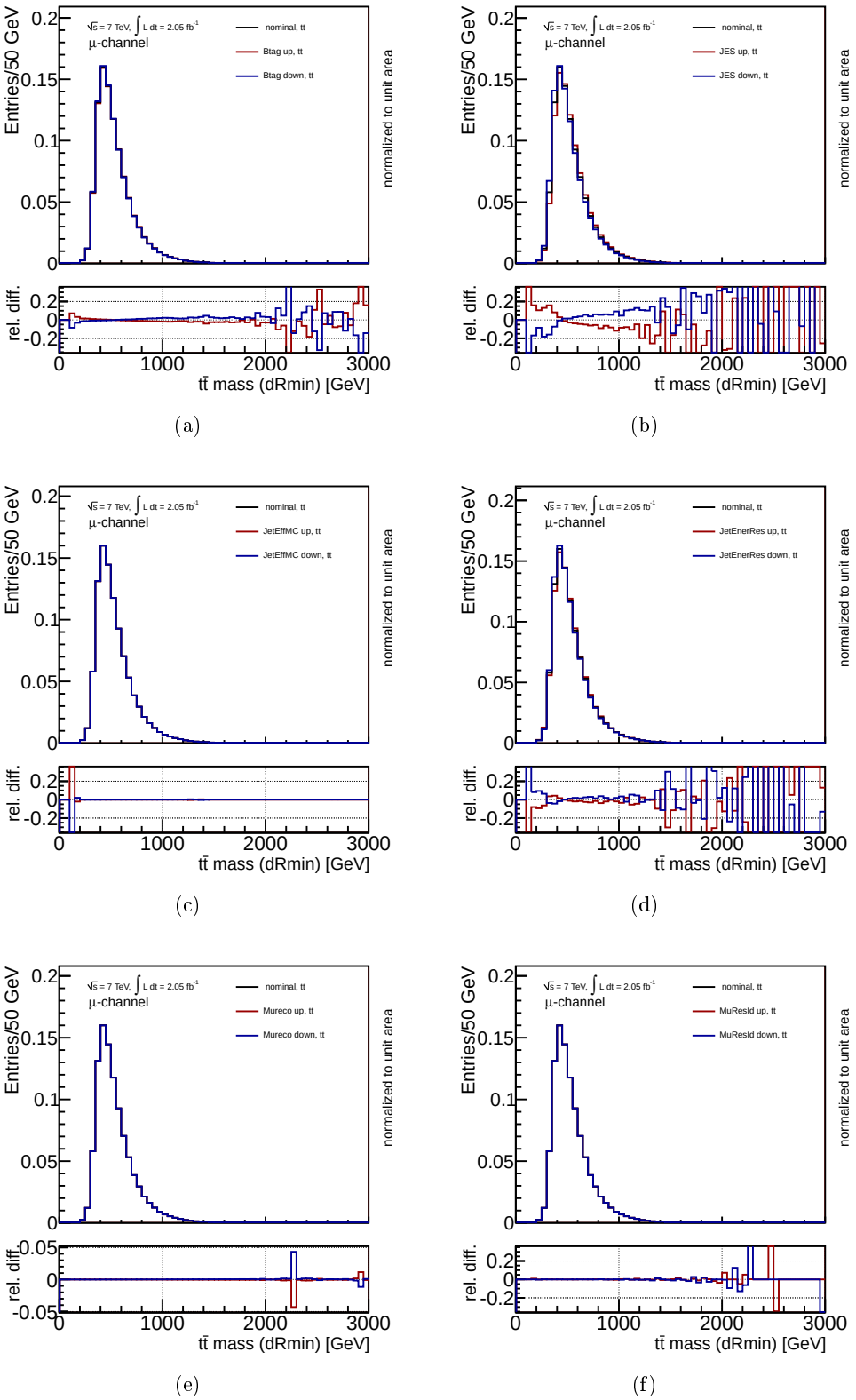


Figure E.6: Impact the (a) b-tagging efficiency, (b) jet energy scale, (c) jet reconstruction efficiency, (d) jet energy resolution, (e) muon reconstruction efficiency and (f) muon p_{T} resolution systematic uncertainties on the $t\bar{t}$ mass spectra for the SM $t\bar{t}$ background in the muon channel [8].

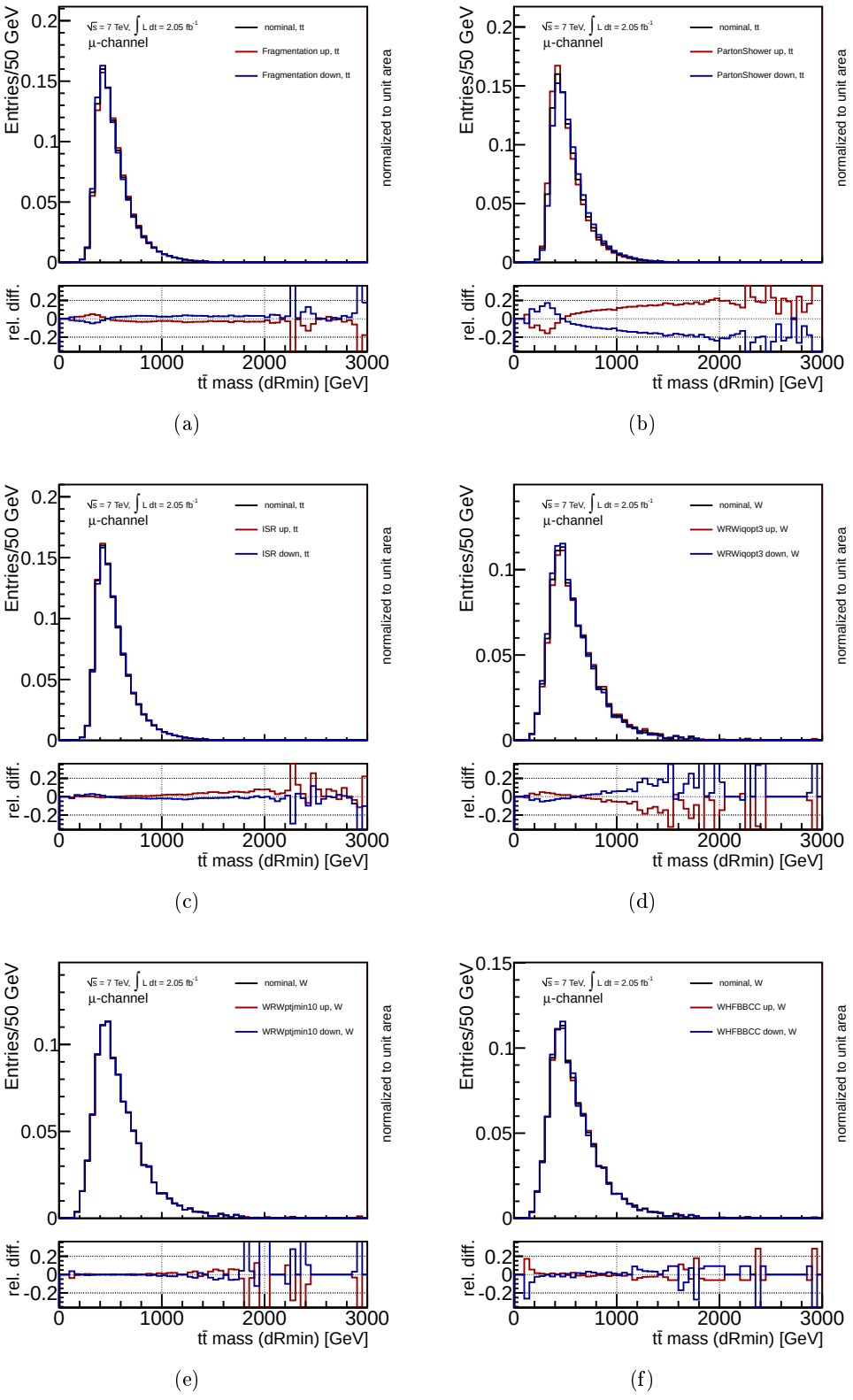


Figure E.7: Impact of the (a) $m_{t\bar{t}}$ shape, (b) parton shower and fragmentation and (c) ISR systematic uncertainties on the $t\bar{t}$ mass spectra for the SM $t\bar{t}$ background in the muon channel. The impact of the (d) W+jets shape (`iqopt3`), (e) W+jets shape (`ptjmin`) and (f) W+jets flavor composition systematic uncertainties on the $t\bar{t}$ mass spectra for the W+jets background in the muon channel [8].

Bibliography

[1] ATLAS Collaboration. Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at $\sqrt{s}$=7 TeV. Technical Report ATLAS-CONF-2010-053, CERN, Geneva, Jul 2010. (Cited on pages ii, vii, 2, 51, 64 and 77.)

[2] E Busato, R Camacho, G Choudalakis, E W Hughes, D Lopez Mateos, K Perez, and A Schwartzman. Validation of the Jet Global Sequential Calibration Using Proton-Proton Collisions at $\sqrt{s} = 7$ TeV and Evaluation of its Associated Systematic Uncertainty. Technical Report ATL-PHYS-INT-2011-060, CERN, Geneva, Jul 2011. (Cited on pages ii, vii, 2, 80, 87, 90 and 91.)

[3] ATLAS Collaboration. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. (arXiv:1112.6426v1), 2011. (Cited on pages ii, vii, 2, 64, 65, 66, 67, 84, 97, 101, 102 and 141.)

[4] Georges Aad et al. A search for ttbar resonances with the ATLAS detector in 2.05 fb$^-$1 of proton-proton collisions at $\sqrt{s} = 7$ TeV. *CERN-PH-EP-2012-122*, 2012. (Cited on pages iii, vii, 147 and 179.)

[5] B Acharya et al. Object selection and calibration, background estimations and MC samples for the Winter 2012 Top Quark analyses with 2011 data. Technical Report ATL-COM-PHYS-2012-224, CERN, Geneva, Feb 2012. (Cited on pages iii, viii, 2, 58 and 67.)

[6] B Acharya et al. Object selection and calibration, background estimations and MC samples for the Autumn 2012 Top Quark analyses with 2011 data. Technical Report ATL-COM-PHYS-2012-1197, CERN, Geneva, Aug 2012. (Cited on pages iii, viii and 2.)

[7] ATLAS Collaboration. A Search for $t\bar{t}$ Resonances in the Lepton Plus Jets Channel in 200 pb^{-1} of pp Collisions at $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2011-087, CERN, Geneva, Jun 2011. (Cited on pages 2, 107 and 108.)

[8] ATLAS Collaboration. A Search for $t\bar{t}$ Resonances in the Lepton Plus Jets Channel using 2.05 fb^{-1} of pp Collisions at $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2012-029, CERN, Geneva, Mar 2012. (Cited on pages 2, 120, 121, 122, 126, 130, 143, 145, 146, 147, 196, 197, 204 and 205.)

[9] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Ludwig, et al. Updated Status of the Global Electroweak Fit and Constraints on New Physics. (arXiv:1107.0975), 2011. (Cited on pages 3, 10 and 15.)

[10] D.H. Perkins. *Introduction To High Energy Physics*. Cambridge University Press, 2000. (Cited on page 3.)

[11] F. Halzen and A.D. Martin. *Quarks and leptons: an introductory course in modern particle physics*. Wiley, 1984. (Cited on page 3.)

[12] G Altarelli. The Standard model of particle physics. (arXiv:hep-ph/0510281v1), 2005. To appear in the 'Encyclopedia of Mathematical Physics, Elsevier'. (Cited on page 3.)

[13] K Nakamura and Particle Data Group. Review of Particle Physics. *Journal of Physics G: Nuclear and Particle Physics*, 37(7A):075021, 2010. (Cited on pages 4, 8, 16, 122 and 134.)

[14] G. Dissertori, I.G. Knowles, and M. Schmelling. *Quantum chromodynamics: high energy experiments and theory*. International series of monographs on physics. Clarendon Press, Oxford, 2003. (Cited on page 6.)

[15] G. P. Salam. Elements of QCD for hadron colliders. (arXiv:1011.5131v2 [hep-ph]), November 2010. (Cited on pages 6 and 47.)

[16] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: Evidence, candidates and constraints. *Phys.Rept.*, 405:279–390, 2005. (Cited on page 9.)

[17] E. Komatsu et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophys.J.Suppl.*, 180:330–376, 2009. (Cited on page 9.)

[18] Y. Ashie et al. Evidence for an Oscillatory Signature in Atmospheric Neutrino Oscillations. *Phys. Rev. Lett.*, 93:101801, Sep 2004. (Cited on page 9.)

[19] R. Barate et al. Search for the standard model Higgs boson at LEP. *Phys.Lett.*, B565:61–75, 2003. (Cited on page 9.)

[20] CDF Collaboration and D0 Collaboration. Combined CDF and D0 Search for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data. (arXiv:hep-ex/0306033v1), 2012. Preliminary results prepared for the Winter 2012 Conferences. (Cited on pages 10 and 11.)

[21] An update to the combined search for the Standard Model Higgs boson with the ATLAS detector at the LHC using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2012-019, CERN, Geneva, Mar 2012. (Cited on page 10.)

[22] CMS Collaboration. Combination of SM, SM4, FP Higgs boson searches. (CMS-PAS-HIG-12-008), Mar 2012. (Cited on page 10.)

[23] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *ArXiv e-prints*, July 2012. (Cited on page 10.)

[24] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *ArXiv e-prints*, (arXiv:1207.7235), July 2012. (Cited on page 10.)

[25] Stephen P. Martin. A Supersymmetry primer. (arXiv:hep-ph/9709356v6), 1997. (Cited on page 10.)

[26] Gian F. Giudice. Beyond the standard model. (arXiv:hep-ph/9605390v1), 1996. (Cited on pages 10 and 12.)

[27] Estia Eichten and Kenneth Lane. Dynamical breaking of weak interaction symmetries. *Physics Letters B*, 90(1–2):125–130, 1980. (Cited on page 12.)

[28] Christopher T Hill. Topcolor: top quark condensation in a gauge extension of the standard model. *PhysLett*, B266(3-4):419–424, 1991. (Cited on page 12.)

[29] Christopher T. Hill. Topcolor assisted technicolor. *Phys.Lett.*, B345:483–489, 1995. (Cited on page 12.)

[30] Robert M. Harris, Christopher T. Hill, and Stephen J. Parke. Cross Section for Topcolor Z' decaying to top-antitop. (arXiv:hep-ph/9911288v1), 1999. (Cited on pages 12, 106, 107 and 110.)

[31] Theodor Kaluza. On the Problem of Unity in Physics. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1921:966–972, 1921. (Cited on page 13.)

[32] O. Klein. Quantum theory and five-dimensional theory of relativity. *Z. Phys.*, 37:895–906, 1926. (Cited on page 13.)

[33] L. Randall and R. Sundrum. A Large mass hierarchy from a small extra dimension. *Phys.Rev.Lett.*, 83:3370–3373, 1999. 9 pages, LaTex Report-no: MIT-CTP-2860, PUPT-1860, BUHEP-99-9. (Cited on page 13.)

[34] B. Lillie, L. Randall, and L. T. Wang. The Bulk RS KK-gluon at the LHC. *JHEP*, 0709:074, 2007. (Cited on pages 13 and 107.)

[35] K. Agashe, A. Belyaev, T. Krupovnickas, G. Perez, and J. Virzi. LHC Signals from Warped Extra Dimensions. *Phys.Rev.*, D77:015003, 2008. (Cited on pages 13 and 107.)

[36] A. Djouadi, G. Moreau, and R. K. Singh. Kaluza-Klein excitations of gauge bosons at the LHC. *Nucl.Phys.*, B797:1–26, 2008. (Cited on page 13.)

[37] N. Arkani-Hamed, S. Dimopoulos, and G.R. Dvali. The Hierarchy problem and new dimensions at a millimeter. *Phys.Lett.*, B429:263–272, 1998. (Cited on page 13.)

[38] Vardan Khachatryan et al. Search for Microscopic Black Hole Signatures at the Large Hadron Collider. *Phys.Lett.*, B697:434–453, 2011. (Cited on page 14.)

[39] Georges Aad et al. (arXiv:1204.4646v1 [hep-ex]), 2012. (Cited on page 14.)

[40] John H. Schwarz. Introduction to superstring theory. (arXiv:hep-ex/0008017v1), 2000. (Cited on page 14.)

[41] F. Abe et al. Observation of top quark production in $\bar{p}p$ collisions. *Phys.Rev.Lett.*, 74:2626–2631, 1995. (Cited on page 14.)

[42] S. Abachi et al. Observation of the top quark. *Phys.Rev.Lett.*, 74:2632–2637, 1995. (Cited on page 14.)

[43] CDF Collaboration and D0 Collaboration. Combination of CDF and D0 results on the mass of the top quark using up to 5.8 fb-1 of data. (arXiv:1107.5255v3 [hep-ex]), 2011. 16 pages, 2 figures. (Cited on page 15.)

[44] Viatcheslav Shary and Pedro Ferreira da Silva. Rencontres de Moriond 2012. La Thuile, Italy, Mar 2012. (Cited on page 15.)

[45] N. Kidonakis. Next-to-next-to-leading soft-gluon corrections for the top quark cross section and transverse momentum distribution. *Phys.Rev.*, D82:114030, 2010. (Cited on page 15.)

[46] N. Kidonakis. Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production. *Phys.Rev.*, D83:091503, 2011. (Cited on pages 15, 109 and 154.)

[47] N. Kidonakis. NNLL resummation for s-channel single top quark production. *Phys.Rev.*, D81:054028, 2010. (Cited on pages 16, 109 and 154.)

[48] N. Kidonakis. Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-. *Phys.Rev.*, D82:054018, 2010. (Cited on pages 16, 109 and 154.)

[49] T. Aaltonen et al. Evidence for a Mass Dependent Forward-Backward Asymmetry in Top Quark Pair Production. *Phys.Rev.*, D83:112003, 2011. 23 pages, 18 figures, submitted to Physical Review D. (Cited on page 18.)

[50] Lyndon Evans and Philip Bryant. LHC Machine. *JINST*, 3(S08001), 2008. (Cited on page 19.)

[51] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3(S08002), 2008. (Cited on page 19.)

[52] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3(S08003), 2008. (Cited on pages 19, 24, 26, 28, 30, 31, 33, 34 and 38.)

[53] R. Adolphi et al. The CMS experiment at the CERN LHC. *JINST*, 3(S08004), 2008. (Cited on page 19.)

[54] A. Alves et al. The LHCb Detector at the LHC. *JINST*, 3(S08005), 2008. (Cited on page 19.)

[55] Xabier Cid Vidal and Ramon Cid. This is a test entry of type @ONLINE: http://www.lhc-closer.es. (Cited on page 20.)

[56] ATLAS Collaboration. Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC. *Eur. Phys. J.*, C71(1630), 2011. (Cited on pages 21, 108, 138 and 153.)

[57] ATLAS collaboration. Updated Luminosity Determination in pp Collisions at $\sqrt{s}$ =7 TeV Using the ATLAS detector. (ATLAS-CONF-2011-011), 2011. (Cited on pages 21, 108 and 153.)

[58] ATLAS Collaboration. dE/dx measurement in the ATLAS Pixel Detector and its use for particle identification. Technical Report ATLAS-CONF-2011-016, CERN, Geneva, Mar 2011. (Cited on page 25.)

[59] R. Wigmans. *Principes et techniques de calorimétrie*. September 1988. (Cited on page 27.)

[60] M. Aleksa et al. Measurement of the ATLAS solenoid magnetic field. *JINST*, 3(P04003), 2008. (Cited on page 36.)

[61] P. Jenni, M. Nordberg, M. Nessi, and K. Jon-And. ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity. *CERN-LHCC*, (2008-004), 2008. (Cited on page 36.)

[62] S. van der Meer. Calibration of the Effective Beam Height in the ISR. *CERN-ISR-PO-68-31*, 1968. (Cited on page 38.)

[63] P. Calafiura, W. Lavrijsen, C. Leggett, M. Marino, and D. Quarrie. The athena control framework in production, new developments and lessons learned. *Computing in High-Energy Physics (CHEP '04)*, pages 456–458, 2005. (Cited on page 41.)

[64] R.N. Cahn and G. Goldhaber. *The experimental foundations of particle physics*. Cambridge University Press, 2009. (Cited on page 42.)

[65] Richard Keith Ellis, William James Stirling, and Bryan R Webber. *QCD and Collider Physics*. Cambridge monographs on particle physics, nuclear physics, and cosmology. Cambridge Univ. Press, Cambridge, 2003. Photography by S. Vascotto. (Cited on pages 42 and 47.)

[66] Richard P. Feynman. Very High-Energy Collisions of Hadrons. *Phys. Rev. Lett.*, 23(24):1415–1417, Dec 1969. (Cited on page 42.)

[67] J C Collins and D E Soper. The Theorems of Perturbative QCD. *Annual Review of Nuclear and Particle Science*, 37(1):383–409, 1987. (Cited on page 43.)

[68] Michelangelo L Mangano and T J Stelzer. Tools for the simulation of hard hadronic collisions. *Annu. Rev. Nucl. Part. Sci.*, 55(CERN-PH-TH-2005-074):555–588, 2005. (Cited on pages 44 and 47.)

[69] Andy Buckley, Jonathan Butterworth, Stefan Gieseke, David Grellscheid, Stefan Hoche, et al. General-purpose event generators for LHC physics. *Phys.Rept.*, 504:145–233, 2011. (Cited on pages 44 and 77.)

[70] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972. (Cited on page 47.)

[71] G. Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298, 1977. (Cited on page 47.)

[72] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977. (Cited on page 47.)

[73] S. Catani, F. Krauss, R. Kuhn, and B.R. Webber. QCD matrix elements + parton showers. *JHEP*, 0111:063, 2001. (Cited on page 47.)

[74] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, and Michele Treccani. Matching matrix elements and shower evolution for top-quark production in hadronic collisions. *JHEP*, 0701:013, 2007. (Cited on page 47.)

[75] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur.Phys.J.*, C63:189–285, 2009. (Cited on page 48.)

[76] B. Andersson, G. Gustafson, and B. Söderberg. A probability measure on parton and string states. *Nuclear Physics B*, 264(0):29–59, 1986. (Cited on page 49.)

[77] B.R. and Webber. A QCD model for jet fragmentation including soft gluon interference. *Nuclear Physics B*, 238(3):492–528, 1984. (Cited on page 49.)

[78] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006. (Cited on pages 50 and 77.)

[79] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al. HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 0101:010, 2001. (Cited on page 50.)

[80] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al. HERWIG 6.5 release note. 2002. (Cited on page 50.)

[81] J.M. Butterworth, Jeffrey R. Forshaw, and M.H. Seymour. Multiparton interactions in photoproduction at HERA. *Z.Phys.*, C72:637–646, 1996. (Cited on page 50.)

[82] M. Bahr, S. Gieseke, M.A. Gigg, D. Grellscheid, K. Hamilton, et al. Herwig++ Physics and Manual. *Eur.Phys.J.*, C58:639–707, 2008. (Cited on pages 50 and 77.)

[83] M. Bähr, S. Gieseke, and M. H. Seymour. Simulation of multiple partonic interactions in Herwig++. *Journal of High Energy Physics*, 7:76, July 2008. (Cited on page 50.)

[84] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307:001, 2003. (Cited on page 50.)

[85] Stefano Frixione and Bryan R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 0206:029, 2002. (Cited on page 51.)

[86] Stefano Frixione, Paolo Nason, and Bryan R. Webber. Matching NLO QCD and parton showers in heavy flavor production. *JHEP*, 0308:007, 2003. (Cited on page 51.)

[87] J. M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nuclear Physics B Proceedings Supplements*, 205:10–15, August 2010. (Cited on pages 51 and 154.)

[88] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer. MadGraph/MadEvent v4: the new web generation. *Journal of High Energy Physics*, 9:28, September 2007. (Cited on page 51.)

[89] Borut Paul Kersevan and Elzbieta Richter-Was. The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5. (TPJU-6-2004), 2004. (Cited on page 51.)

[90] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 0711:070, 2007. (Cited on page 51.)

[91] John Allison. Facilities and Methods: Geant4 – A Simulation Toolkit. *Nuclear Physics News*, 17(2):20–24, 2007. (Cited on pages 51, 78, 109 and 153.)

[92] ATLAS Collaboration. The ATLAS Simulation Infraestructure. *Eur.Phys.J.*, C(70):823–874, 2008. (Cited on pages 51, 78, 109 and 153.)

[93] ATLAS Collaboration. Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at $\sqrt{s}$=900 GeV. Technical Report ATLAS-CONF-2010-016, CERN, Geneva, Jul 2010. (Cited on pages 51, 64 and 77.)

[94] ATLAS Collaboration. Mapping the material in the ATLAS Inner Detector using secondary hadronic interactions in 7 TeV collisions. Technical Report ATLAS-CONF-2010-058, CERN, Geneva, Jul 2010. (Cited on page 51.)

[95] ATLAS Collaboration. Alignment Performance of the ATLAS Inner Detector Tracking System in 7 TeV proton-proton collisions at the LHC. Technical Report ATLAS-CONF-2010-067, CERN, Geneva, Jul 2010. (Cited on page 51.)

[96] ATLAS Collaboration. ATLAS Calorimeter Response to Single Isolated Hadrons and Estimation of the Calorimeter Jet Scale Uncertainty. Technical Report ATLAS-CONF-2010-052, CERN, Geneva, Jul 2010. (Cited on page 51.)

[97] ATLAS Collaboration. Muon Reconstruction Performance. Technical Report ATLAS-CONF-2010-064, CERN, Geneva, Jul 2010. (Cited on page 51.)

[98] G. Aad et al. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics. (arXiv:0901.0512v4 [hep-ex]), 2009. (Cited on page 53.)

[99] T Cornelissen, M Elsing, S Fleischmann, W Liebig, E Moyse, and A Salzburger. Concepts, Design and Implementation of the ATLAS New Tracking (NEWT). Technical Report ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar 2007. (Cited on page 54.)

[100] ATLAS Collaboration. Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment. Technical Report ATLAS-CONF-2012-042, CERN, Geneva, Mar 2012. (Cited on pages 54 and 56.)

[101] ATLAS Collaboration. Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment. Technical Report ATLAS-CONF-2010-069, CERN, Geneva, Jul 2010. (Cited on page 55.)

[102] Georges Aad et al. Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data. *Eur.Phys.J.*, C72:1909, 2012. (Cited on pages 56, 57, 58 and 63.)

[103] M. Aharrouche et al. Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 614(3):400–432, 2010. (Cited on pages 56 and 63.)

[104] E. Abat, J. M. Abdallah, T. N. Addy, P. Adragna, M. Aharrouche, A. Ahmad, T. P. A. Akesson, M. Aleksa, C. Alexa, K. Anderson, and et al. Combined performance studies for electrons at the 2004 ATLAS combined test-beam. *Journal of Instrumentation*, 5:11006, November 2010. (Cited on pages 56 and 63.)

[105] ATLAS Collaboration. Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector. Technical Report ATLAS-CONF-2011-063, CERN, Geneva, Apr 2011. (Cited on pages 59 and 60.)

[106] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 0804:063, 2008. (Cited on pages 62 and 79.)

[107] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Phys.Rev.*, D48:3160–3166, 1993. (Cited on page 62.)

[108] M. Wobisch and T. Wengler. Hadronization corrections to jet cross-sections in deep inelastic scattering. (arXiv:hep-ph/9907280v1), 1998. (Cited on page 62.)

[109] ATLAS Collaboration. Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s}$ = 7 TeV. Technical Report ATLAS-CONF-2010-054, CERN, Geneva, Jul 2010. (Cited on pages 63 and 141.)

[110] P. Adragna et al. Testbeam studies of production modules of the ATLAS Tile Calorimeter. *Nucl. Instrum. Meth.*, A(606):362–394, 2009. (Cited on page 63.)

[111] J. Pinfold et al. Performance of the ATLAS liquid argon endcap calorimeter in the pseudorapidity region in beam tests. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 593(3):324–342, 2008. (Cited on page 63.)

[112] G Romeo, A Schwartzman, R Piegaia, T Carli, and R Teuscher. Jet Energy Resolution from In-situ Techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s} = 7$ TeV. Technical Report ATL-COM-PHYS-2011-240, CERN, Geneva, Mar 2011. (Cited on pages 65, 67, 82, 85 and 88.)

[113] ATLAS Collaboration. In-situ jet energy scale and jet shape corrections for multiple interactions in the first ATLAS data at the LHC. (ATLAS-CONF-2011-030), February 2011. (Cited on page 65.)

[114] P Loch, J A Backus Mayes, G Zevi Della Porta, C Sandoval, S Menke, F Vives, A Schwartzman, and P Manning. Pile-up corrections for jets from proton-proton collisions at $\sqrt{s} = 7$ TeV in ATLAS in 2011. Technical Report ATL-COM-PHYS-2012-349, CERN, Geneva, Mar 2012. (Cited on page 65.)

[115] D Lopez Mateos, E W Hughes, and A Schwartzman. A Simple p_T- and η-Dependent Monte Carlo-Based Jet Calibration. Technical Report ATL-PHYS-INT-2009-077, CERN, Geneva, Aug 2009. (Cited on pages 66 and 73.)

[116] ATLAS Collaboration. Update on the energy scale systematic uncertainty for jets jet produced in proton-proton collisions at $\sqrt{s} = 7$ TeV measured with the ATLAS detector. (ATLAS-CONF-2011-007), February 2011. (Cited on pages 67, 97, 101 and 141.)

[117] ATLAS Collaboration. Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data. Technical Report ATLAS-CONF-2011-102, CERN, Geneva, Jul 2011. (Cited on page 67.)

[118] ATLAS Collaboration. Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector. Technical Report ATLAS-CONF-2012-043, CERN, Geneva, Mar 2012. (Cited on page 69.)

[119] ATLAS Collaboration. Calibrating the b-Tag Efficiency and Mistag Rate in 35 pb^{-1} of Data with the ATLAS Detector. Technical Report ATLAS-CONF-2011-089, CERN, Geneva, Jun 2011. (Cited on page 69.)

[120] ATLAS Collaboration. Measurement of the Mistag Rate with 5 fb^1 of Data Collected by the ATLAS Detector. Technical Report ATLAS-CONF-2012-040, CERN, Geneva, Mar 2012. (Cited on page 69.)

[121] ATLAS Collaboration. b-jet tagging calibration on c-jets containing D^{*+} mesons. Technical Report ATLAS-CONF-2012-039, CERN, Geneva, Mar 2012. (Cited on page 69.)

[122] Georges Aad et al. Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS. *Eur.Phys.J.*, C72:1844, 2012. (Cited on pages 69, 70 and 141.)

[123] ATLAS Collaboration. Light-quark and Gluon Jets in ATLAS. Technical Report ATLAS-CONF-2011-053, CERN, Geneva, Apr 2011. (Cited on pages 72, 84 and 86.)

[124] D Lopez Mateos, E W Hughes, and A Schwartzman. A Sequential Multi-Variate Jet Calibration Based On Global Properties of the Jet Structure. Technical Report ATL-COM-PHYS-2010-058, CERN, Geneva, Jan 2010. (Cited on pages 72, 73, 79 and 89.)

[125] D Lopez Mateos, E W Hughes, and A Schwartzman. Jet Energy Resolution Improvement After Calibration Using Longitudinal Calorimeter Segmentation in ATLAS. Technical Report ATL-PHYS-INT-2009-051, CERN, Geneva, Apr 2009. (Cited on pages 72, 74 and 79.)

[126] ATLAS Collaboration. Charged particle multiplicities in pp interactions for track pT>100 MeV at $\sqrt{s} = 0.9$ and 7 TeV measured with the ATLAS detector at the LHC. Technical Report ATLAS-CONF-2010-046, CERN, Geneva, Jul 2010. (Cited on page 77.)

[127] Peter Zeiler Skands. Tuning Monte Carlo Generators: The Perugia Tunes. *Phys.Rev.*, D82:074018, 2010. (Cited on page 77.)

[128] Georges Aad et al. Performance of the ATLAS Trigger System in 2010. *Eur.Phys.J.*, C72:1849, 2012. (Cited on page 78.)

[129] Matteo Cacciari and Gavin P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Phys.Lett.*, B641:57–61, 2006. (Cited on page 79.)

[130] ATLAS Collaboration. Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2010-038, CERN, Geneva, Jul 2010. (Cited on page 79.)

[131] D Lopez Mateos, E W Hughes, and A Schwartzman. Reduction of the Flavor Dependence of the Jet Energy Scale Using a Correction that Depends on the Transverse Structure of the Jet. Technical Report ATL-PHYS-INT-2010-046, CERN, Geneva, Apr 2010. (Cited on page 85.)

[132] ATLAS Collaboration. In-situ Pseudo-rapidity Inter-calibration to Evaluate Jet Energy Scale Uncertainty and Calorimeter Performance in the Forward Region. *ATLAS-CONF 2010-055*, 2010. (Cited on page 88.)

[133] ATLAS Collaboration. Determination of the ATLAS jet energy measurement uncertainty using photon-jet events in proton-proton collisions at sqrts = 7 TeV. Technical Report ATLAS-CONF-2011-031, CERN, Geneva, Mar 2011. (Cited on pages 99 and 100.)

[134] ATLAS Collaboration. Prospects for top anti-top resonance searches using early ATLAS data. Technical Report ATL-PHYS-PUB-2010-008, CERN, Geneva, Jul 2010. (Cited on pages 106 and 114.)

[135] A Arce et al. A Search for $t\bar{t}$ Resonances in Lepton+Jets Events with Highly Boosted Top Quarks collected in 2 fb^{-1} of pp Collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. Technical Report ATL-COM-PHYS-2012-306, CERN, Geneva, Mar 2012. (Cited on pages 106 and 107.)

[136] V.M. Abazov et al. Search for $t\bar{t}$ resonances in the lepton plus jets final state in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys.Lett.*, B668:98–104, 2008. (Cited on pages 106 and 107.)

[137] T. Aaltonen et al. A Search for resonant production of $t\bar{t}$ pairs in 4.8 fb^{-1} of integrated luminosity of $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV. *Phys.Rev.*, D84:072004, 2011. (Cited on pages 106, 107 and 108.)

[138] T. Aaltonen et al. Search for resonant $t\bar{t}$ production in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV. *Phys.Rev.Lett.*, 100:231801, 2008. 7 pages, 3 figures. submitted to Phys. Rev. Lett. (Cited on page 107.)

[139] T. Aaltonen et al. Limits on the production of narrow $t\bar{t}$ resonances in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV. *Phys.Rev.*, D77:051102, 2008. (Cited on page 107.)

[140] T. Aaltonen et al. Search for New Color-Octet Vector Particle Decaying to ttbar in ppbar Collisions at $\sqrt{s} = 1.96$ TeV. *Phys.Lett.*, B691:183–190, 2010. (Cited on page 107.)

[141] T. Aaltonen et al. Search for resonant production of $t\bar{t}$ decaying to jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys.Rev.*, D84:072003, 2011. (Cited on pages 107 and 108.)

[142] Victor Mukhamedovich Abazov et al. Search for a Narrow $t\bar{t}$ Resonance in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. *Phys.Rev.*, D85:051101, 2012. 7 pages, 3 figures, submitted to Physical Review Letters. (Cited on pages 107 and 108.)

[143] S. Swedish and on behalf of the ATLAS Collaboration. A search for ttbar resonances in the dilepton channel in 1.04/fb of pp collisions at sqrt(s) = 7 TeV with the ATLAS detector. 2012. (Cited on pages 107 and 108.)

[144] CMS Collaboration. Search for Z' bosons decaying to top-antitop pairs in dilepton+jets final state in pp collisions at $\sqrt{s}$. (CMS-PAS-TOP-11-010), Apr 2012. (Cited on pages 107 and 108.)

[145] CMS Collaboration. A Search for Resonances in Semileptonic Top Pair Production. (CMS-PAS-TOP-11-009), 2012. (Cited on pages 107 and 108.)

[146] CMS Collaboration. Search for high-mass resonances decaying to ttbar in the electrons+jets channel. (CMS-PAS-EXO-11-092), Feb 2012. (Cited on pages 107 and 108.)

[147] CMS Collaboration. Search for heavy narrow resonances decaying to ttbar in the muon+jets channel. (CMS-PAS-EXO-11-055), Aug 2011. (Cited on pages 107 and 108.)

[148] CMS Collaboration. Search for BSM $t\bar{t}$Production in the Boosted All-Hadronic Final State. (CMS-PAS-EXO-11-006), Jul 2011. (Cited on pages 107 and 108.)

[149] Pavel M. Nadolsky, Hung-Liang Lai, Qing-Hong Cao, Joey Huston, Jon Pumplin, et al. Implications of CTEQ global analysis for collider observables. *Phys.Rev.*, D78:013004, 2008. (Cited on page 109.)

[150] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann. HATHOR - HAdronic Top and Heavy quarks crOss section calculatoR. *Computer Physics Communications*, 182:1034–1046, April 2011. (Cited on pages 109 and 154.)

[151] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, Pavel M. Nadolsky, et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 0207:012, 2002. (Cited on page 109.)

[152] J Butterworth, E Dobson, U Klein, B Mellado Garcia, T Nunnemann, J Qian, D Rebuzzi, and R Tanaka. Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV. Technical Report ATL-COM-PHYS-2010-695, CERN, Geneva, Aug 2010. (Cited on pages 109 and 154.)

[153] A. Sherstnev and R.S. Thorne. Parton Distributions for LO Generators. *Eur.Phys.J.*, C55:553–575, 2008. (Cited on page 109.)

[154] ATLAS Collaboration. Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data. Technical Report ATLAS-CONF-2012-020, CERN, Geneva, Mar 2012. (Cited on page 113.)

[155] Georges Aad et al. Measurement of the charge asymmetry in top quark pair production in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector. (arXiv:1203.4211v1 [hep-ex]), 2012. (Cited on page 116.)

[156] Georges Aad et al. Measurement of the cross section for the production of a W boson in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys.Lett.*, B707:418–437, 2012. (Cited on page 116.)

[157] K Becker, A Cortes Gonzalez, V Dao, F Derue, K Gellerstedtf, D Hirschbuehl, J Howarth, H Khandanyan, F Kohn, T M Liss, M A Owen, M Pinamonti, E Shabalina, P Sturm, A Succurro, T Theveneaux-Pelzerd, W Wagner, W H Bell, and J Sjolin. Mis-identified lepton backgrounds in top quark pair production studies for EPS 2011 analyses. Technical Report ATL-COM-PHYS-2011-768, CERN, Geneva, Jun 2011. (Cited on page 119.)

[158] Thorsten Chwalek. *Messung der W-Boson-Helizitätsanteile in Top-Quark-Zerfällen mit dem CDF II Experiment und Studien zu einer frühen Messung des ttbar-Wirkungsquerschnitts mit dem CMS Experiment. oai:cds.cern.ch:1416031.* PhD thesis, Karlsruhe, U., Karlsruhe, 2010. Presented 12 Feb 2010. (Cited on page 125.)

[159] S. Moch and P. Uwer. Theoretical status and prospects for top-quark pair production at hadron colliders. *Phys. Rev. D*, 78:034003, Aug 2008. (Cited on page 138.)

[160] Georges Aad et al. Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV. *Eur.Phys.J.*, C71:1577, 2011. 30 pages plus author list (50 pages total), 9 figures, 11 tables, CERN-PH number and final journal added. (Cited on page 138.)

[161] G. Choudalakis. On hypothesis testing, trials factor, hypertests and the BumpHunter. *arXiv:1101.0390v2 [physics.data-an]*, January 2011. (Cited on page 142.)

[162] I. Bertram et al. A Recipe for the construction of confidence limits. FERMILAB-TM-2104. (Cited on page 144.)

[163] H.B. Prosper, Supriya Jain, and R. Schwienhorst. Statistical methods implemented in the package top statistics. *D0 internal note*, 5817, 2008. (Cited on page 144.)

[164] D W Miller, A Schwartzman, and D Su. Jet-Vertex Association Algorithm. Technical Report ATL-COM-PHYS-2008-008, CERN, Geneva, Jan 2008. (Cited on pages 150 and 151.)

[165] D W Miller, A Schwartzman, and D Su. Pile-up jet energy scale corrections using the jet-vertex fraction method. Technical Report ATL-PHYS-INT-2009-090, CERN, Geneva, Sep 2009. (Cited on page 151.)

[166] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, et al. New parton distributions for collider physics. *Phys.Rev.*, D82:074024, 2010. (Cited on page 154.)

[167] ATLAS collaboration. New ATLAS event generator tunes to 2010 data. Technical Report ATL-PHYS-PUB-2011-008, CERN, Geneva, Apr 2011. (Cited on page 154.)

[168] K Becker, A Cortes Gonzalez, V Dao, F Derue, K Gellerstedtf, D Hirschbuehl, J Howarth, H Khandanyan, F Kohn, T M Liss, M A Owen, M Pinamonti, E Shabalina, P Sturm, A Succurro, T Theveneaux-Pelzerd, W Wagner, W H Bell, and J Sjolin. Mis-identified lepton backgrounds in top quark pair production studies for EPS 2011 analyses. Technical Report ATL-COM-PHYS-2011-768, CERN, Geneva, Jun 2011. (Cited on page 160.)